

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 27, 2003, 16:41:03 ; Search time 82 Seconds  
(without alignments)  
11982.833 Million cell updates/sec

Title: US-09-830-433a-7

Perfect score: 3204

Sequence: 1 atggaacgaaccccaacctt.....gcgtaggctaccggtctga 3204

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA.\*  
1: /cgn2.6/ptodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2.6/ptodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2.6/ptodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2.6/ptodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2.6/ptodata/1/ina/PCTUS\_COMB.seq.\*  
6: /cgn2.6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49.6	1.5	913	4	US-08-818-112-16
2	49.6	1.5	913	4	US-08-818-111-16
3	49.6	1.5	913	4	US-09-056-556-16
4	49.6	1.5	913	4	US-09-072-596-16
5	49.6	1.5	4403765	4	US-09-103-840A-2
6	49.6	1.5	4411529	4	US-09-103-840A-1
7	47.8	1.5	7218	1	US-08-232-463-14
8	45.2	1.4	4403765	4	US-09-103-840A-2
9	45.2	1.4	4411529	4	US-09-103-840A-1
10	41.4	1.3	2949	4	US-09-412-554A-3
11	39.4	1.2	2367	4	US-09-056-556-201
12	39.4	1.2	2367	4	US-09-072-596-196
13	39.2	1.2	625	4	US-08-998-416-1112
14	37.8	1.2	329	4	US-09-056-556-168
15	37.8	1.2	329	4	US-09-072-596-163
16	37.6	1.2	494	4	US-09-056-556-176
17	37.6	1.2	494	4	US-09-072-596-171
18	37.4	1.2	3572	4	US-09-575-574-3
19	37	1.2	610	4	US-09-072-596-293
20	36.8	1.1	954	2	US-08-336-198C-2
21	36.6	1.1	6530	2	US-08-146-930-1
22	36.6	1.1	6530	3	US-08-458-240-1
23	36.6	1.1	6530	5	PCT-US93-03993-1
24	36.4	1.1	1631	4	US-08-895-590-43
25	36.2	1.1	1280	4	US-09-060-750-4
26	36.2	1.1	1399	1	US-08-471-033-24
27	36.2	1.1	1399	2	US-08-471-044-24

28	36.2	1.1	1399	2	US-08-463-483A-24	Sequence 24, Appl
29	36.2	1.1	1399	2	US-08-471-046A-24	Sequence 24, Appl
30	36.2	1.1	1399	2	US-08-470-566B-24	Sequence 24, Appl
31	36.2	1.1	1399	2	US-08-469-334-24	Sequence 24, Appl
32	36.2	1.1	1399	3	US-09-300-529-24	Sequence 3, Appl
c 33	36	1.1	935	1	US-08-162-475A-3	Sequence 26, Appl
34	35.6	1.1	4162	3	US-08-459-595A-26	Sequence 26, Appl
35	35.6	1.1	4162	3	US-08-459-595A-26	Sequence 26, Appl
36	35.6	1.1	4162	3	US-08-459-504B-26	Sequence 26, Appl
37	35.6	1.1	4162	3	US-08-459-444-26	Sequence 26, Appl
38	35.6	1.1	4162	4	US-09-547-422-26	Sequence 26, Appl
39	35.6	1.1	4165	1	US-07-951-715A-26	Sequence 26, Appl
40	35.6	1.1	4689	4	US-09-105-537-34	Sequence 34, Appl
41	35.6	1.1	36778	4	US-09-105-537-5	Sequence 5, Appl
42	35.6	1.1	38506	3	US-09-320-878-19	Sequence 19, Appl
43	35.2	1.1	538	4	US-09-056-556-180	Sequence 180, App
44	35.2	1.1	538	4	US-09-072-596-175	Sequence 175, App
c 45	35.2	1.1	991	4	US-09-344-529-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1  
US-08-818-112-16  
; Sequence 16, Application US/08818112  
; Patent No. 6290969  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Campos-Neto, Antonio  
; APPLICANT: Houghton, Raymond  
; APPLICANT: Twardzik, Daniel R.  
; APPLICANT: Twardzik, Daniel R.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY  
; NUMBER OF SEQUENCES: 153  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED AND BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/818,112  
; FILING DATE: 13-MAR-1997  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Maki, David J.  
; REGISTRATION NUMBER: 31,392  
; REFERENCE/DOCKET NUMBER: 210121.411C6  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 913 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-818-112-16

Query Match 1.5%; Score 49.6; DB 4; Length 913;  
Best Local Similarity 47.7%; Pred. No. 0.00048;  
Matches 136; Conservative 1; Mismatches 148; Indels 0; Gaps 0;



Db 725 GCTGCTGAAGCCGTTAGCGCGGTTCCGCGSGGTTCCGCGGGTGGCGCCNTGGCGCCGCGG 784

QY 2245 GCAGCGGACGCGGTACAGCATGCGAATGCCGCGAGCGGTGTACGCATCTTCAACAGCTCTC 2304

Db 785 CCGCGGTTGGCGTACAGCACACCCCGGTGGCGCGTTGGCGCCATTGCGCGCCATTGCC 844

QY 2305 GCGGCTACGCTATGCGGACAGTACCGCGCCCATGCCGATATG 2349

Db 845 GCGGTTGCGCGCATTCGCGCGGTTCCGCGCGCACCGCGGNTTG 889

RESULT 4

US-09-072-596-16

; Sequence 16, Application US/09072596

; Patent No. 6458366

GENERAL INFORMATION:

; APPLICANT: Reed, Steven G.

; APPLICANT: Shelly, Yasir A.W.

; APPLICANT: Dillon, Davin C.

; APPLICANT: Campos-Neto, Antonia

; APPLICANT: Houghton, Raymond

; APPLICANT: Vedvick, Thomas S.

; APPLICANT: Twardzik, Daniel R.

; APPLICANT: Lodes, Michael J.

; APPLICANT: Hendrickson, Ronald C.

; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF

NUMBER OF SEQUENCES: 350 TUBERCULOSIS

CORRESPONDENCE ADDRESS:

; ADDRESSER: SEED AND BERRY LLP

; STREET: 6300 Columbia Center, 701 Fifth Avenue

; CITY: Seattle

; STATE: Washington

; COUNTRY: USA

; ZIP: 98104-7092

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/072,596

; FILING DATE: 05-MAY-1998

CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Maki, David J.

; REGISTRATION NUMBER: 31,392

; REFERENCE/DOCKET NUMBER: 210121.417C9

TELECOMMUNICATION INFORMATION:

; TELEPHONE: (206) 622-4900

; TELEFAX: (206) 682-6031

; INFORMATION FOR SEQ ID NO: 16:

SEQUENCE CHARACTERISTICS:

; LENGTH: 913 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

US-09-072-596-16

Query Match 1.5%; Score 49.6; DB 4; Length 913;

Best Local Similarity 47.7%; Pred. No. 0.00048;

Matches 136; Conservative 1; Mismatches 148; Indels 0; Gaps 0;

QY 2065 GCGGCACATTCGCGCGCGCGTCTGAAACAGCGCGGTAGAACAGGCGCGCAGCAATCTG 2124

Db 605 GCGCGGAGNGCTGCGCGCGCGCGCGCAAGCCCAAAAGCCCGGGTTGCCACCGCG 664

QY 2125 GAAACCTGATGCTGCACTGGATGCTCCGAATCATCCGAACACACCGAGAGCGTTGAA 2184

Db 665 CCGCGCGGACCCACCGTCCCGCGCATCCCGCGTTCGCGCGCGTCCCGCGCATTTGGT 724

QY 2185 ACTGCGGCGCGCGCACAGATATGCGGGGATCCGCCCTACGCGCGCACTTTCGCG 2244

Db 725 GCTGCTGAAGCCGTTAGCGCGGTTCCGCGSGGTTCCGCGGGTGGCGCCNTGGCGCGCGG 784

QY 2245 GCAGCGGACGCGGTACAGCATGCGAATGCCGCGAGCGGTGTACGCATCTTCAACAGCTCTC 2304

Db 785 CCGCGGTTGGCGTACAGCACACCCCGGTGGCGCGTTGGCGCCATTGCGCGCCATTGCC 844

QY 2305 GCGGCTACGCTATGCGGACAGTACCGCGCCCATGCCGATATG 2349

Db 845 GCGGTTGCGCGCATTCGCGCGGTTCCGCGCGCACCGCGGNTTG 889

RESULT 5

US-09-103-840A-2

; Sequence 2, Application US/09103840A

; Patent No. 6294328

GENERAL INFORMATION:

; APPLICANT: FLEISCHMAN, Robert D.

; APPLICANT: WHITE, Owen R.

; APPLICANT: FRASER, Claire M.

; APPLICANT: VENTER, John C.

; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM

; TITLE OF INVENTION: TUBERCULOSIS

; FILE REFERENCE: 24366-20007 00

; CURRENT APPLICATION NUMBER: US/09/103,840A

; CURRENT FILING DATE: 1998-06-24

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2

; LENGTH: 4403765

; TYPE: DNA

; ORGANISM: Mycobacterium tuberculosis

; FEATURE:

; OTHER INFORMATION: CDC 1551

; OTHER INFORMATION: "n" bases at various positions throughout the sequence

; OTHER INFORMATION: represent a, t, c or g

US-09-103-840A-2

Query Match 1.5%; Score 49.6; DB 4; Length 4403765;

Best Local Similarity 44.5%; Pred. No. 0.041;

Matches 196; Conservative 0; Mismatches 244; Indels 0; Gaps 0;

QY 2076 GCGCGCGCGCGTCTGAAACACGCGGTAGAACAGGCGCGCAAGTCTGAAAACCTGAT 2135

Db 336988 CGCTGCGCGCGTGGCGCGCGCGCGCGCGGAGCGCGGTAGCGCGCGGTGC 337047

QY 2136 GGTCAACTGGATGCTCGGAATCATCCGCAACACCGAGACGCTTGAAACTCGCGCGC 2195

Db 337048 GCGCGTGGCGCGTGGCGCGCGCAAGTGGCGCGCTGTGCGCGCAGCACCGCAGTGC 337107

QY 2196 CGACCGCACAGATATGCGGGCATCCGCCCTACGGCGCAACTTTCGCGCAGCGCAGC 2255

Db 337108 GCGCGCGCGGAACAGCGCGCGCTGCCCCCGCGCGCGTCCGCGCGCGAGCGTGC 337167

QY 2256 CGTACAGCATGCGAATGCGCGCGCGCGGTGTACGCATCTTCAACAGTCTCGCGCTAC 2315

Db 337168 CCGCGCGGTGTGCGCGCGCGCGCGCAAGAGCAATCCGTTCCCGCGGTCCCGCATTC 337227

QY 2316 CTATGCGCAGTACCGCGCGCCCATGCCGATATGAGGGAGCGCGGTGAAAGCGGTATC 2375

Db 337228 CAAACCGCGCGCGCGCGCGCGCGCGCGCGGTGGCGACACCGCACCGGTACCGCGCTC 337287

QY 2376 GGACGGTTGGACCAACAGCTACGGGTCTGCGCGTCTATCGCGCAAAACCAACAGGAC 2435

Db 337288 GCGCGTGGCGCGCGCACCGATAAGTTTGGGAGAGGGCGGCTGCGCGCGGTCCCTG 337347

QY 2436 TGAAGCTGGGAACAGGGCGGTGTTGAGGCAAAATGCGCGGCGAGTACCCAAACCGTC 2495

Db 337348 GCGCACCGAGAACAGCGCGGTACCGCGCGCGCGCGCGCGCGCGGTGTCCAGGC 337407

QY 2496 CATTCGCGGAAACCGCGG 2515

Db 337408 CAAACCGCGGTGCGCGCG 337427

## RESULT 6

US-09-103-840A-1  
; Sequence 1, Application US/09103840A  
; Patent No. 6294328  
; GENERAL INFORMATION:  
; APPLICANT: FLEISCHMAN, Robert D.  
; APPLICANT: WHITE, Owen R.  
; APPLICANT: FRASER, Claire M.  
; APPLICANT: VENTER, John C.  
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
; TITLE OF INVENTION: TUBERCULOSIS  
; FILE REFERENCE: 24366-20007.00  
; CURRENT APPLICATION NUMBER: US/09/103.840A  
; CURRENT FILING DATE: 1998-06-24  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 4411529  
; TYPE: DNA  
; ORGANISM: Mycobacterium tuberculosis  
; OTHER INFORMATION: H37RV  
US-09-103-840A-1

Query Match 1.5%; Score 49.6; DB 4; Length 4411529;  
Best Local Similarity 44.1%; Pred. No. 0.041;  
Matches 208; Conservative 0; Mismatches 264; Indels 0; Gaps 0;  
  
Qy 2044 GCGGACGAGCTGCTCGGACGCGCACATTCGCGCGCGCGCGCTCTGAACACGCGCGTA 2103  
Db 336867 GCGGACGCGCGGCTCGGCGCGCGCACATTCGCGCGCGCGCGCTCTGAACACGCGCGCG 336926  
  
Qy 2104 GAACAGGCGCGGAGCAATCTGGAACATCTGATGTCGAACTGATGCTCGGAATCATCC 2163  
Db 336927 GGAGCGCGGAGGCGAGTAGGCGCGCGTTCGCGCGCGTTCGCGCGCGAGGT 336986  
  
Qy 2164 GGAACACCGGAGAGCGTTGAACATCTCGGCGCGCGCGCGCGATATCGCGGCGATCGCG 2223  
Db 336987 GCGCGCGTGTGCG 337046  
  
Qy 2224 CCCTACGCGGCAACTTTCG 2283  
Db 337047 CCG 337106  
  
Qy 2284 GTACGCATCTTCAACAGTCTCGCGCGCTACCGCTCTATGCGGACGATGCGCGCGCGCGCG 2343  
Db 337107 CAATCGTTCCTCCCGCGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 337166  
  
Qy 2344 GATATGAGGAGCGCGCGCTGAAGCGCTATGCGGACGCGGTTGGACCAACGCTACGCGT 2403  
Db 337167 GTTGGCGAACACCG 337226  
  
Qy 2404 CTGCGCGTATCGCGCGCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2463  
Db 337227 TTGGGAGAGGCGCGCGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 337286  
  
Qy 2464 GCGAAATACG 2515  
Db 337287 GCG 337338

## RESULT 7

US-08-232-463-14/c  
; Sequence 14, Application US/08232463  
; Patent No. 5670367  
; GENERAL INFORMATION:  
; APPLICANT: DORNER, F.  
; APPLICANT: SCHEIFLINGER, F.  
; APPLICANT: FALKNER, F. G.  
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 1800 Diagonal Road, Suite 500

; CITY: Alexandria  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22313-0299  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA: US/08/232.463  
; APPLICATION NUMBER: US/08/232.463  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/935.313  
; FILING DATE:  
; APPLICATION NUMBER: EP 91 114 300.6  
; FILING DATE: 26-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)836-9300  
; TELEFAX: (703)683-4109  
; TELEX: 899149  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7218 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; CLONE: pTZgpl-F15  
US-08-232-463-14

Query Match 1.5%; Score 47.8; DB 1; Length 7218;

Best Local Similarity 6.0%; Pred. No. 0.0045;

Matches 25; Conservative 216; Mismatches 178; Indels 0; Gaps 0;

Qy 432 CATATCCTTTCCGNACTGTATGGCAGAAAGAACACACGCGCTATACGAAATACAAAA 491  
Db 1476 CTTATCTATGCAAGTAGTAAAGAGATAGAAATTTGGTACRRRRRRRRRRRRRRRR 1417  
  
Qy 492 CTATACGCGGTATATGCGGAAGAGCGCTGAAGACGAGCGGTAAGACATTAAAGC 551  
Db 1416 RRR 1357  
  
Qy 552 TTCTTTTCGACGATGAGCGCGTTATAGAGACTGAAGCAAGCGGATATCCGCCAGT 611  
Db 1356 RRR 1297  
  
Qy 612 AAAGAAATCGCACATCGATGTGCTCCCATATTATTGCGCGCGCTTCGTCGACGG 671  
Db 1296 RRR 1237  
  
Qy 672 CAGACCTCAGCGGTATTGCGCGCGCTGCGACGCTACACATAATACGATGATGG 731  
Db 1236 RRR 1177  
  
Qy 732 AACCAAGAACAAATAATGTCTGACGCGCATCCGCAATCGATGGTCAAGTGGCGAAG 791  
Db 1176 RRR 1117  
  
Qy 792 TGGCGTGGCGATCGCAATACAGTTTGGACACACATCGAGGCGGCGACTGCGGAC 850  
Db 1116 RRR 1058

## RESULT 8

US-09-103-840A-2/c  
; Sequence 2, Application US/09103840A  
; Patent No. 6294328





```
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/056,556
; FILING DATE: 07-APR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.457
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 201:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2367 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-056-556-201

Query Match      1.2%; Score 39.4; DB 4; Length 2367;
Best Local Similarity 50.8%; Pred. No. 0.54;
Matches 94; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

Qy 2164 GCAACACCCGAGAGCGTTGAAACTGCGCGCGCCGACGATATATCGCGGGCATCCGC 2223
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 119 GCCACCGGAGCGCGGTTGCGCGCTACCGCCGCGGACTCCACCGGCCACC 178

Qy 2224 CCCTACGGCGCAACTTTCGGCGCAGCGCGGTACAGCATGCGAATGCCCGACCGT 2283
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 179 GACTCCGCGGTGCGACCGTTGCGCGCTTCCGCGATCAACATGCGCGTGGCGCCACCCTT 238

Qy 2284 GTACGCATCTTCAACAGTCTCGCGCTACCGTCTATGCGGACAGTACCGCGCCATGCC 2343
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 239 GCCACCCAGCCAGCGGCTCGCCACCACCGCGGACCAACAGGAGCTGCGCGGAGCC 298

Qy 2344 GATAT 2348
    ||
Db 299 ACCAT 303

RESULT 12
US-09-072-596-196
; Sequence 196, Application US/09072596
; Patent No. 6458366
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Campos-Neto, Antonia
; APPLICANT: Houghton, Raymond
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Twardzik, Daniel R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Hendrickson, Ronald C.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF TUBERCULOSIS
; NUMBER OF SEQUENCES: 350
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/072,596
; FILING DATE: 05-MAY-1998
; CLASSIFICATION:
```

```
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.417C9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 196:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2367 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-072-596-196

Query Match      1.2%; Score 39.4; DB 4; Length 2367;
Best Local Similarity 50.8%; Pred. No. 0.54;
Matches 94; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

Qy 2164 GCAACACCCGAGAGCGTTGAAACTGCGCGCGCCGACGATATATCGCGGGCATCCGC 2223
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 119 GCCACCGGAGCGCGGTTGCGCGCTACCGCCGCGGACTCCACCGGCCACC 178

Qy 2224 CCCTACGGCGCAACTTTCGGCGCAGCGCGGTACAGCATGCGAATGCCCGACCGT 2283
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 179 GACTCCGCGGTGCGACCGTTGCGCGCTTCCGCGATCAACATGCGCGTGGCGCCACCCTT 238

Qy 2284 GTACGCATCTTCAACAGTCTCGCGCTACCGTCTATGCGGACAGTACCGCGCCATGCC 2343
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 239 GCCACCCAGCCAGCGGCTCGCCACCACCGCGGACCAACAGGAGCTGCGCGGAGCC 298

Qy 2344 GATAT 2348
    ||
Db 299 ACCAT 303

RESULT 13
US-08-998-416-1112
; Sequence 1112, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Reibischung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPII
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwalis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
```

REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-541-8587  
TELEFAX: 919-541-8689  
INFORMATION FOR SEQ ID NO: 1112:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 625 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE: PAG16770P  
ORGANISM: PAG16770P  
US-08-998-416-1112

Query Match 1.2%; Score 39.2; DB 4; Length 625;  
Best Local Similarity 48.6%; Pred. No. 0.31;  
Matches 107; Conservative 0; Mismatches 113; Indels 0; Gaps 0;  
QY 2197 GACCGCACAGATATGCGGGCATCGGCCCTACGGGCGCAACTTTCCGCGCAGCGGCAGCC 2256  
DB 290 GCCGGCACCTACGTGCGCTCGCCACTCAACTGGCGCTCTCCGTGCGGCGCGGTGCGCC 349  
QY 2257 GTACAGCATCGGAATGCGCGCGCGAGGTGTACGCATCTTCAACAGTCTTCGCGGTACCGTC 2316  
DB 350 CTGGAGCACGTGGACAGCTGGAGCGCGCGCGGTGTCACGCGTCCGCGAGCGGC 409  
QY 2317 TATGCGCACAGTACCGCGCCCATCGCGATATGCGGAGCGCGGCTGAAGCGGTATCG 2376  
DB 410 TCGTACGTCTATGATGTATGTCTCGGCCCTGCAGCGGGCTTCGCGCGGCCCTTTGCGCGC 469  
QY 2377 GACGGGTGGACCAACACGCTACGGGTCTCGCGGTCTATCG 2416  
DB 470 CATCCGGTGTGCCACACCTTTGCGGCGCTCGCGCGCGCG 509

RESULT 14  
US-09-056-556-168/c  
Sequence 168, Application US/09056556  
Patent No. 6350456  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Dillon, Davin C.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND  
NUMBER OF SEQUENCES: 241  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/056,556  
FILING DATE: 07-APR-1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.457  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 168:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 329 base pairs  
TYPE: nucleic acid

STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-056-556-168  
Query Match 1.2%; Score 37.8; DB 4; Length 329;  
Best Local Similarity 49.3%; Pred. No. 0.54;  
Matches 99; Conservative 0; Mismatches 102; Indels 0; Gaps 0;  
QY 2150 CCTCCGAATCATCGCAACACCCGAGACGCTTGAACACTGCGCGCGCCGACGACGATA 2209  
DB 244 CCGCGGCACCTCGCTTGAAGGCTGTTCGCGGGCGCGCGCGCGGTACCGCCA 185  
QY 2210 TGCCGGGCATCCGCCCTACGCGCGCAACTTTCCGCGCAGCGCGCGGTACAGCATGCGA 2269  
DB 184 ACACCGCGCGCGCGCGGTCCCGCGGCCCGCGCGCGCTTTGGCGCGCCAGCAGCTG 125  
QY 2270 ATGCGCGCGAGGTGTACGATCTTCAACAGTCTTCGCGGTACCGTGTATGCGCGACAGTA 2329  
DB 124 ATCAGCGTTA 65  
QY 2330 CCGCGCGCCATCGCGATATGC 2350  
DB 64 CCACCATCACCGCGGTATCC 44

RESULT 15  
US-09-072-596-163/c  
Sequence 163, Application US/09072596  
Patent No. 6458366  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Dillon, Davin C.  
APPLICANT: Campos-Neto, Antonia  
APPLICANT: Houghton, Raymond  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Twardzik, Daniel R.  
APPLICANT: Lodes, Michael J.  
APPLICANT: Hendrickson, Ronald C.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF  
NUMBER OF SEQUENCES: 350  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/072,596  
FILING DATE: 05-MAY-1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.417C9  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 163:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 329 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-072-596-163  
Query Match 1.2%; Score 37.8; DB 4; Length 329;



GENERAL INFORMATION:  
APPLICANT: Yu, Su-May  
TITLE OF INVENTION: PLANT SEEDLING AND EMBRYO PROMOTER  
FILE REFERENCE: 08919-047001  
CURRENT APPLICATION NUMBER: US/09/575,574  
CURRENT FILING DATE: 2000-05-22  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 3572  
TYPE: DNA  
ORGANISM: Oryza sativa  
US-09-575-574-3

Query Match 1.2%; Score 37.4; DB 4; Length 3572;  
Best Local Similarity 48.0%; Pred. No. 2.4;  
Matches 107; Conservative 0; Mismatches 116; Indels 0; Gaps 0;  
QY 2149 GCTCTCGAATCATCCGCAACACCCGAGAGCGTTGAAACTGCGCGCCGCGACGACAGAT 2208  
DB 2535 GCGCGCGTAGCGCGCCGACCGGTGTCACCGTAGCGCGCGCGCGCGCGCGCGCGCC 2476  
QY 2209 ATGCGCGGATCGCGCCCTACGCGCAACTTTCCGCGCAGCGCGCGCGTACGACATGCG 2268  
DB 2475 GCGGTAGCC 2416  
QY 2269 AATGCG 2328  
DB 2415 TCCTCGGTAGCG 2356  
QY 2329 ACCG 2371  
DB 2355 CCGCGCTCG 2313

RESULT 19  
US-09-072-596-293  
Sequence 293, Application US/09072596  
Patent No. 6458366  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Dillion, Davin C.  
APPLICANT: Campos-Neto, Antonia  
APPLICANT: Houghton, Raymond  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Twardzik, Daniel R.  
APPLICANT: Lodes, Michael J.  
APPLICANT: Hendrickson, Ronald C.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF TUBERCULOSIS  
NUMBER OF SEQUENCES: 350  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED AND BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/072,596  
FILING DATE: 05-MAY-1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.417C9  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 293:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 610 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-09-072-596-293  
Query Match 1.2%; Score 37; DB 4; Length 610;  
Best Local Similarity 49.2%; Pred. No. 1.2;  
Matches 97; Conservative 0; Mismatches 100; Indels 0; Gaps 0;  
QY 2180 TTGAAGTGGCG 2239  
DB 30 TTGCTCTCGGTGCG 89  
QY 2240 TCCG 2299  
DB 90 GGTCTTCTACGTGCG 149  
QY 2300 GTCTCGCGGTACCGTCTATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2359  
DB 150 GCCCGCGCGTACCG 209  
QY 2360 GGCTGAAGCGGTATCG 2376  
DB 210 GCCTGGTTGACGTATTG 226  
RESULT 20  
US-08-336-198C-2/c  
Sequence 2, Application US/08336198C  
Patent No. 5866382  
GENERAL INFORMATION:  
APPLICANT: Hallborn, Johan  
APPLICANT: Penttila, Merja  
APPLICANT: Ojamo, Heikki  
APPLICANT: Keranen, Sirkka  
APPLICANT: Hahn-Hagerdal, Barbel  
APPLICANT: Walfridsson, Mats  
APPLICANT: Airaksinen, Ulla  
TITLE OF INVENTION: XYLOSE UTILIZATION BY RECOMBINANT YEASTS  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Birch, Stewart, Kolasch & Birch  
STREET: 301 N. Washington St.  
CITY: Falls Church  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22046-3487  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/336,198C  
FILING DATE: 03-NOV-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Murphy Jr., Gerald M.  
REGISTRATION NUMBER: 28,977  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-205-8000  
TELEFAX: 703-205-8050  
TELEX: 248345  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 954 base pairs  
TYPE: nucleic acid

STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Pichia stipitis  
STRAIN: CBS-6054  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..954  
OTHER INFORMATION: /standard\_name= "xylose reductase"  
PUBLICATION INFORMATION: FI 901771  
DOCUMENT NUMBER: FI 901771  
FILING DATE: 06-APR-1990  
US-08-336-198C-2

Query Match 1.1%; Score 36.8; DB 2; Length 954;  
Best Local Similarity 50.6%; Pred. No. 1.8;  
Matches 89; Conservative 0; Mismatches 87; Indels 0; Gaps 0;  
QY 2640 CTATCTCAAGGCTTCTCTCTACGGAGCTTACAAACAGCATCAGCCGAGCAGCCGG 2699  
DB 193 CTTCTGTAATGCTTCTTGACACCGGACCACTAATTTTCTGTTGGGTAACTTCGG 134  
QY 2700 TCGGACGAACATCGGAGGAGCGGTCAACGGCAGCGTGTGATGCGAGCTGGGCGCACTGGG 2759  
DB 133 CACCGTCGAACAATCTGTAACCGGTCTTGTATAGCAGCGGTAGATCTGTTTCAAGAACAGGTGT 74  
QY 2760 CGGTGTCAACGTTTCGTTTGGCGAACGGAGATTGACGGTCGAGGCGGTCTGC 2815  
DB 73 CGAGTCGAGTTTCCAAACAGCCGAAACCGGCTGGCATGTGCTGAACAGAGTTTC 18

## RESULT 21

US-08-146-930-1/c  
; Sequence 1, Application US/08146930  
; Patent No. 5958764

GENERAL INFORMATION:  
; APPLICANT: Roop, Dennis R.  
; APPLICANT: Rothnagel, Joseph A.  
; APPLICANT: Greenhalgh, David A.  
; TITLE OF INVENTION: SPECIFIC EXPRESSION VECTORS  
; TITLE OF INVENTION: AND METHODS OF USE  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LYON & LYON  
; STREET: 611 West Sixth Street  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90017  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: IBM MS-DOS (Version 5.0)  
; SOFTWARE: Wordperfect (Version 5.1)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/146,930  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; PRIOR APPLICATION DATA: including application  
; PRIOR APPLICATION DATA: described below:  
; APPLICATION NUMBER: 07/876,286  
; FILING DATE: April 30, 1992  
; APPLICATION NUMBER: No. 5958764 yet assigned (204/132)  
; FILING DATE: October 29, 1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warburg, Richard J.  
; REGISTRATION NUMBER: 32,327  
; REFERENCE/DOCKET NUMBER: 204/152  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (213) 489-1600

TELEFAX: (213) 955-0440  
TELEX: 67-3510  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6530 bases  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-08-146-930-1

Query Match 1.1%; Score 36.6; DB 2; Length 6530;  
Best Local Similarity 45.8%; Pred. No. 5.5;  
Matches 126; Conservative 0; Mismatches 149; Indels 0; Gaps 0;

QY 2063 CAGCGGCACATTCGCGCCCGCGGTCTGAACACGCGGTAGAACAGGCGCGCACAATC 2122  
DB 3142 CGGGAGTAGTCTTGACCGCCCGCCACAGAGCGCGCACGCTGCCACAGCTGGAGCCACCCCG 3083  
QY 2123 TGGAAACCTGATGTCGAACCTGGATGCTCCGAATCATCCGCAACACCCGAGAGCGGTG 2182  
DB 3082 GAGTAGCTGCCGCCCGCGAGCTGGAGCGCGCGCGCTCCGAGTAGCCACCTCCGCGAG 3023  
QY 2183 AAATGCGCGCGCGCGCGACGCGACAGATATCGCGGGCATCCGCCCTTACGCGCAACTTCC 2242  
DB 3022 CTGGAGCGCGCGCGCGCGAGTAGGACCGCGCGAGTAGAGCCACCGCGCTCCG 2963  
QY 2243 CGCGAGCGCGCGCGCTACAGCATGCGAATGCGCGCGAGGTGTACGCATCTTCAACAGTC 2302  
DB 2962 GAGTAGCGCGCGCGCGAGTAGAGCCACCGCGCGCTCCGAGTAGTTCGACGCCCGCACCG 2903  
QY 2303 TCGCGCTACCGTCTATGCGCGAGTAGTACCGCGCGC 2337  
DB 2902 GAGCGCGCTCCACCGCGCGAGCTGGAACCACTCC 2868

## RESULT 22

US-08-458-240-1/c  
; Sequence 1, Application US/08458240  
; Patent No. 6143727

GENERAL INFORMATION:  
; APPLICANT: Roop, Dennis R.  
; APPLICANT: Rothnagel, Joseph A.  
; APPLICANT: Greenhalgh, David A.  
; TITLE OF INVENTION: SPECIFIC EXPRESSION VECTORS  
; TITLE OF INVENTION: AND METHODS OF USE  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LYON & LYON  
; STREET: 611 West Sixth Street  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90017  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: IBM MS-DOS (Version 5.0)  
; SOFTWARE: Wordperfect (Version 5.1)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/458,240  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/146,930  
; FILING DATE:  
; APPLICATION NUMBER: 07/876,286  
; FILING DATE: April 30, 1992  
; APPLICATION NUMBER: No. 6143727 yet assigned (204/132)  
; FILING DATE: October 29, 1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warburg, Richard J.  
; REGISTRATION NUMBER: 32,327

REFERENCE/DOCKET NUMBER: 204/152  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6530 bases  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-458-240-1

Query Match 1.1%; Score 36.6; DB 3; Length 6530;  
Best Local Similarity 45.8%; Pred. No. 5.5;  
Matches 126; Conservative 0; Mismatches 149; Indels 0; Gaps 0;  
QY 2063 CAGCGGCACATTCGCGCGCGCGTCTGAACACACCGCTAGAACAGCGCGGCGGAGCAATC 2122  
DB 3142 CCGGAGTACTTGACGCCGCCACACAGAGCGCCACCGCTCCACAGCTGGAGCCACCCCG 3083  
QY 2123 TGAACAACTGTGTCGAACCTGGATGCTCCGAATCATCCGGAATCATCCGGAACACCGGAGACGGTTG 2182  
DB 3082 GAGTAGCTGCGCGCGCGAGCTGGAGCGCGCGCTCCGGAGTAGCCACCTCCGCGAG 3023  
QY 2183 AAACCTGCGCGCGCGACGACAGATGCGGGGCATCCGCCCTACGGCGCAACTTCC 2242  
DB 3022 CTGAGCGCGCGCGCGCGAGTAGCCACCGCGGAGCTAGAGCCACCGCGCTCCG 2963  
QY 2243 GCGCAGCGCGCGCTACAGCATGCGAATGCGCGCGCGAGCGGTGTACGCATCTTCAACAGTC 2302  
DB 2962 GAGTAGCGCGCGCGAGCTAGAGCCACCGCGCGCTCCGGAGTACTTGAGCGCCCCACCG 2903  
QY 2303 TCGCCGCTACCGTCTATGCGGACAGTACCGCGCC 2337  
DB 2902 GAGCGCGCTCCACCGCGCGAGTGGAAACCACTCC 2868

RESULT 23  
PCT-US93-03993-1/C  
Sequence 1, Application PC/TUS9303993  
GENERAL INFORMATION:  
APPLICANT: Roop, Dennis R.  
APPLICANT: Rothnagel, Joseph A.  
APPLICANT: Greenhalgh, David A.  
TITLE OF INVENTION: CONSTITUTIVE AND INDUCIBLE EPIDERMAL  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fulbright & Jaworski  
STREET: 1301 McKinney, Suite 5100  
CITY: Houston  
STATE: Texas  
COUNTRY: U.S.A.  
ZIP: 77010-3095  
COMPUTER READABLE FORM:  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/03993  
FILING DATE: 1993/04/28  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Paul, Thomas D.  
REGISTRATION NUMBER: 32,714  
REFERENCE/DOCKET NUMBER: D-5405  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 713/651-5325  
TELEFAX: 713/651-5246  
TELEX: 762829

INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6530 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
PCT-US93-03993-1

Query Match 1.1%; Score 36.6; DB 5; Length 6530;  
Best Local Similarity 45.8%; Pred. No. 5.5;  
Matches 126; Conservative 0; Mismatches 149; Indels 0; Gaps 0;  
QY 2063 CAGCGGCACATTCGCGCGCGCGTCTGAACACACCGCTAGAACAGCGCGGCGGAGCAATC 2122  
DB 3142 CCGGAGTACTTGACGCCGCCACACAGAGCGCCACCGCTCCACAGCTGGAGCCACCCCG 3083  
QY 2123 TGAACAACTGTGTCGAACCTGGATGCTCCGAATCATCCGGAACACCGGAGACGGTTG 2182  
DB 3082 GAGTAGCTGCGCGCGCGAGCTGGAGCGCGCGCTCCGGAGTAGCCACCTCCGCGAG 3023  
QY 2183 AAACCTGCGCGCGCGACGACAGATGCGGGGCATCCGCCCTACGGCGCAACTTCC 2242  
DB 3022 CTGAGCGCGCGCGCGAGTAGCCACCGCGGAGCTAGAGCCACCGCGCTCCG 2963  
QY 2243 GCGCAGCGCGCGCTACAGCATGCGAATGCGCGCGCGAGCGGTGTACGCATCTTCAACAGTC 2302  
DB 2962 GAGTAGCGCGCGCGAGCTAGAGCCACCGCGCGCTCCGGAGTACTTGAGCGCCCCACCG 2903  
QY 2303 TCGCCGCTACCGTCTATGCGGACAGTACCGCGCC 2337  
DB 2902 GAGCGCGCTCCACCGCGCGAGTGGAAACCACTCC 2868

RESULT 24  
US-08-895-590-43  
Sequence 43, Application US/088955590  
Patent No. 6207410  
GENERAL INFORMATION:  
APPLICANT: Hall, Linda M.  
APPLICANT: Ren, Dejian  
APPLICANT: Zheng, Wei  
APPLICANT: Dubald, Manuel Marcel Paul  
TITLE OF INVENTION: Genes Encoding an Insect Calcium Channel  
NUMBER OF SEQUENCES: 101  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, LLP  
STREET: 699 Prince Street  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22314-3187  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/895,590  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/374,888  
FILING DATE: 19-JAN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: McGowan, Malcolm M.  
REGISTRATION NUMBER: 39,300  
REFERENCE/DOCKET NUMBER: 022650-263  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-836-6620  
TELEFAX: 703-836-2021







```
;
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/037,057
; FILING DATE: 25-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Spruill, W. Murray
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: CGC 1695/CIP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8615
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1399 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..1386
; OTHER INFORMATION: /note= "Maize optimized DNA
; OTHER INFORMATION: sequence for VIP2A(a) protein from AB78"
US-08-433-483A-24

Query Match 1.1%; Score 36.2; DB 2; Length 1399;
Best Local Similarity 45.1%; Pred. No. 3.2;
Matches 134; Conservative 0; Mismatches 163; Indels 0; Gaps 0;

QY 2466 CAAATGCGCGCAGTACCCAAACCGTCGGCATTCGCCGGAACCGCGGAAATACGAC 2525
DB 339 CAGCATGCGCGCAGCTTCGAGGACGAGATCAAGGACCTGAAGGAGATCGACAAGATGTT 398
QY 2526 AGCAGCGCGCAGTACCCAAACCGTCGGCATTCGCCGGAACCGCGGAAATACGAC 2585
DB 399 CGACAAGACCAACCTGAGCAACAGCATATCACCTACAGACAGTGGAGCCACCACCAT 458
QY 2586 AACCGACAGCATTAAGTCTGTTTTCAGGAGATAGCGGACGATGCGGCGGATATCGGCTATCT 2645
DB 459 CGGCTTCAACAAGAGCGCTGACGCGAGGCAACACCATCAACAGCAGCGCATGGCCAGTT 518
QY 2646 CAAAGCGCTGTTCTCTACGGACGCTACAAAAACAGCATACGCCCGCAGCAGCGGTGCGGA 2705
DB 519 CAAGGAGCAGTTTCCTGGACCGGAGCATCAAGTTCGACAGCTACCTGGACACCCACTGAC 578
QY 2706 CGAACATCGGAAGCAGCGCTCAACGGCAGCGTGTGAGCTGGCGGCGGCGG 2762
DB 579 CGCCAGCAGGTGAGCAGCAGGAGCGCGTGTATCTGAAGGTGACCGTCCCGCGCG 635

RESULT 29
US-08-471-046A-24
; Sequence 24, Application US/08471046A
; Patent No. 5866326
; GENERAL INFORMATION:
; APPLICANT: Warren, Gregory W
; APPLICANT: Kozziel, Michael G
; APPLICANT: Mullins, Martha A
; APPLICANT: Nye, Gordon J
; APPLICANT: Carr, Brian
; APPLICANT: Desai, Nalini M
; APPLICANT: Kostichka, N. Kristy
; APPLICANT: Duck, Nicholas B
; APPLICANT: Estruch, Juan J
; TITLE OF INVENTION: Method For Isolating Vegetative Insecticidal
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 5866326artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA: US/08/471,046A
; APPLICATION NUMBER: US/08/471,046A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/463,483
; FILING DATE: 05-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/314,594
; FILING DATE: 09-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/218,018
; FILING DATE: 23-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/037,057
; FILING DATE: 25-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: CGC1695/CIP3/DIV8 - SOLV4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1399 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..1386
; OTHER INFORMATION: /note= "Maize optimized DNA
; OTHER INFORMATION: sequence for VIP2A(a) protein from AB78"
US-08-471-046A-24

Query Match 1.1%; Score 36.2; DB 2; Length 1399;
Best Local Similarity 45.1%; Pred. No. 3.2;
Matches 134; Conservative 0; Mismatches 163; Indels 0; Gaps 0;

QY 2466 CAAATGCGCGCAGTACCCAAACCGTCGGCATTCGCCGGAACCGCGGAAATACGAC 2525
DB 339 CAGCATGCGCGCAGCTTCGAGGACGAGATCAAGGACCTGAAGGAGATCGACAAGATGTT 398
QY 2526 AGCAGCGCGCAGTACCCAAACCGTCGGCATTCGCCGGAACCGCGGAAATACGAC 2585
DB 399 CGACAAGACCAACCTGAGCAACAGCATATCACCTACAGACAGTGGAGCCACCACCAT 458
QY 2586 AACCGACAGCATTAAGTCTGTTTTCAGGAGATAGCGGACGATGCGGCGGATATCGGCTATCT 2645
DB 459 CGGCTTCAACAAGAGCGCTGACGCGAGGCAACACCATCAACAGCAGCGCATGGCCAGTT 518
QY 2646 CAAAGCGCTGTTCTCTACGGACGCTACAAAAACAGCATACGCCCGCAGCAGCGGTGCGGA 2705
DB 519 CAAGGAGCAGTTTCCTGGACCGGAGCATCAAGTTCGACAGCTACCTGGACACCCACTGAC 578
QY 2706 CGAACATCGGAAGCAGCGCTCAACGGCAGCGTGTGAGCTGGCGGCGGCGG 2762
DB 579 CGCCAGCAGGTGAGCAGCAGGAGCGCGTGTATCTGAAGGTGACCGTCCCGCGCG 635

RESULT 30
US-08-470-566B-24
; Sequence 24, Application US/08470566B
; Patent No. 5872212
; GENERAL INFORMATION:
; APPLICANT: Warren, Gregory W
```

```

; APPLICANT: Koziel, Michael G
; APPLICANT: Mullins, Martha A
; APPLICANT: Nye, Gordon J
; APPLICANT: Carr, Brian
; APPLICANT: Desai, Nalini M
; APPLICANT: Kostichka, N. Kristy
; APPLICANT: Duck, Nicholas B
; APPLICANT: Estruch, Juan J
; TITLE OF INVENTION: No. 5872212el Pesticidal Proteins and Strains
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5872212artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,566B
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/463,483
; FILING DATE: 05-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/314,594
; FILING DATE: 09-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/218,018
; FILING DATE: 23-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/037,057
; FILING DATE: 25-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: CGC1695/CIP3/DIV4 - SQLv4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1399 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..1386
; OTHER INFORMATION: /note= "Maize optimized DNA
; OTHER INFORMATION: sequence for VIP2A(a) protein from AB78"
; US-08-470-566B-24
;
; Query Match 1.1%; Score 36.2; DB 2; Length 1399;
; Best Local Similarity 45.1%; Pred. No. 3.2;
; Matches 134; Conservative 0; Mismatches 163; Indels 0; Gaps 0;
;
; QY 2466 CAAATGCGGCGGAGTACCAACCGTCGGCATTGGCGGAAACCGCGGAAATAGGAC 2525
; ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
; Db 339 CACATGGCGCGGCGGCTTCGAGGACGAGATCAAGGACCTGAAGGAGATCGACAAAGATGT 398
; ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
; QY 2526 AGCAGCGCGGCGGAGTGGCATGGACACACATGAGCGGAAACAGTCCAAATGCAA 2585
; ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
; Db 399 CGACAGACCACTGAGCNACAGCATCATCCTTACAAGAACGTGGAGCCCAACCCAT 458
; ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
; QY 2586 AACCGACAGCATTTAGTCTGTTTTCAGGCATACGGCAGTACGGCGGATATCGCGTATCT 2645
; ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 459 CGGCTTCAACAAGAGCCTGACCGGAGGCAACACCATCAACAGCGACGCCATGGCCCGATT 518
;
; QY 2646 CAAAGGCGCTGTTCTCTACGAGCGCTACAAAACAGCATCAGCGCGACCGGTGCGGA 2705
; ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
; Db 519 CAAGGACGAGTTCTTGAGCGCGGACATCAAGTTGAGACGCTACCTGGACACCCACCTGAC 578
; ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
; QY 2706 CGAACATGCGGAGGAGCGCTCAACGGCAGCGCTGATGATGCTGGGCGCACTGGGCGG 2762
; ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
; Db 579 CGCCGACGAGTGAGCAGCAAGGAGCGCGTGATCCTGAAGTGACCTGCCCGACCGG 635
;
; RESULT 31
; US-08-469-334-24
; Sequence 24, Application US/08469334
; Patent No. 5990383
; GENERAL INFORMATION:
; APPLICANT: Warren, Gregory W
; APPLICANT: Koziel, Michael G
; APPLICANT: Mullins, Martha A
; APPLICANT: Nye, Gordon J
; APPLICANT: Carr, Brian
; APPLICANT: Desai, Nalini M
; APPLICANT: Kostichka, N. Kristy
; APPLICANT: Duck, Nicholas B
; APPLICANT: Estruch, Juan J
; TITLE OF INVENTION: No. 5990383el Pesticidal Proteins and Strains
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,334
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/463,483
; FILING DATE:
; APPLICATION NUMBER: US 08/314,594
; FILING DATE: 09-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/218,018
; FILING DATE: 23-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/037,057
; FILING DATE: 25-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Spruill, W. Murray
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: CGC 1695/CIP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8615
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1399 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..1386
; OTHER INFORMATION: /note= "Maize optimized DNA
; OTHER INFORMATION: sequence for VIP2A(a) protein from AB78"

```



```
; ATTORNEY/AGENT INFORMATION:
; NAME: White, Paul E. Jr.
; REGISTRATION NUMBER: 32011
; REFERENCE/DOCKET NUMBER: PEW/3122/204351
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 935 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; MOLECULE TYPE: cDNA to mRNA
; ORIGINAL SOURCE: Lycopersicon chilense
; US-08-162-475A-3

Query Match 1.1%; Score 36; DB 1; Length 935;
Best Local Similarity 55.6%; Pred. No. 2.9;
Matches 69; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 1443 TAGGAAAGTACATCCGATATTCCTTCCGTAACGACATTTTCAGGCACGGCGG 1502
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 316 TACAAAGCAATATCCCTCCAGTGAATGTTCTGCACCTACCAACCATTTGTTTCATGAGATG 257
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1503 CCTGATCAAAAAAGCGCGCAGCCAACTGCAACTGCACGGCAACAAACACCTATACGGGCAA 1562
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 256 TTGTACCGAAAAGGACGACATTCCTTCTTACGTGTCAGTATCATCACCAGTAGTACCAA 197
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1563 AACG 1566
    |||||
Db 196 AACG 193

RESULT 34
US-08-459-448A-26
; Sequence 26, Application US/08459448A
; Patent No. 5859336
; GENERAL INFORMATION:
; APPLICANT: Koziel, Michael G.
; APPLICANT: Desai, Nalini M.
; APPLICANT: Lewis, Kelly S.
; APPLICANT: Kramer, Vance C.
; APPLICANT: Warren, Gregory W.
; APPLICANT: Evola, Stephen V.
; APPLICANT: Crossland, Lyle D.
; APPLICANT: Wright, Martha S.
; APPLICANT: Merlin, Ellis J.
; APPLICANT: Launis, Karen L.
; APPLICANT: Rothstein, Steven J.
; APPLICANT: Bowman, Cindy G.
; APPLICANT: Dawson, John L.
; APPLICANT: Dunder, Erik M.
; APPLICANT: Pace, Gary M.
; APPLICANT: Suttie, Janet L.
; TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
; TITLE OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5859336artis Corporation
; STREET: Patent & Trademark Dept., 520 White Plains
; STREET: Rd., POB 2005
; CITY: Tarrytown
; STATE: New York
; COUNTRY: USA
; ZIP: 10591-9005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,448A

; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/951,715
; FILING DATE: 25-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/772,027
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Pace, Gary M.
; REGISTRATION NUMBER: 40403
; REFERENCE/DOCKET NUMBER: CGC 1577/CIP/DIV4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)541-8582
; TELEFAX: (919)541-8689
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4162 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1418..1427
; OTHER INFORMATION: /note= "start of mRNA"
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1481..2366
; FEATURE:
; NAME/KEY: intron
; LOCATION: 2367..2451
; FEATURE:
; NAME/KEY: exon
; LOCATION: 2452..2602
; FEATURE:
; NAME/KEY: intron
; LOCATION: 2603..2690
; FEATURE:
; NAME/KEY: exon
; LOCATION: 2691..2804
; FEATURE:
; NAME/KEY: intron
; LOCATION: 2805..2906
; FEATURE:
; NAME/KEY: exon
; LOCATION: 2907..3075
; FEATURE:
; NAME/KEY: intron
; LOCATION: 3076..3177
; FEATURE:
; NAME/KEY: exon
; LOCATION: 3178..3304
; FEATURE:
; NAME/KEY: intron
; LOCATION: 3305..3398
; FEATURE:
; NAME/KEY: exon
; LOCATION: 3399..3498
; FEATURE:
; NAME/KEY: intron
; LOCATION: 3499..3713
; FEATURE:
; NAME/KEY: exon
; LOCATION: 3714..3811
; FEATURE:
; NAME/KEY: promoter
; LOCATION: 1..1477
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /partial
; OTHER INFORMATION: /function= "pollen-specific promoter region"
; OTHER INFORMATION: /evidence= EXPERIMENTAL
```





```

; LOCATION: 3399..3498
; FEATURE:
; NAME/KEY: Intron
; LOCATION: 3499..3713
; FEATURE:
; NAME/KEY: exon
; LOCATION: 3714..3811
; SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-08-459-444-26

Query Match      1.1%; Score 35.6; DB 3; Length 4162;
Best Local Similarity 54.6%; Pred. No. 8.3;
Matches 71; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 38 AACCGGCTGCATGGCGTTAGCTGTTGCAACAACACATTTCTGCTGCTTAGCGCGCGCGC 97
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1581 AACGGGGTGTCTGCTCGCGGTGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGC 1640
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 98 GCGGCACTTCTGGCGCCGAGTTCAATCAGCGCGCACCGGTATCGGCAGCAGCAACAGCAG 157
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1641 GCGGCAGCAGCAAGCGCGGCTCACCCACGCGCGCGCAGGCCAGCTCCGCGCAGCAAC 1700
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 158 CAACAACAGC 167
    | | | | |
Db 1701 CGGCGCGGC 1710

RESULT 38
US-09-547-422-26
; Sequence 26, Application US/09547422
; Patent No. 6320100
; GENERAL INFORMATION:
; APPLICANT: Koziel, Michael G.
;           Desai, Nalini M.
;           Lewis, Kelly S.
;           Kramer, Vance C.
;           Warren, Gregory W.
;           Evola, Stephen V.
;           Crossland, Lyle D.
;           Wright, Martha S.
;           Merlin, Ellis J.
;           Launis, Karen L.
; TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
;                   INSECTICIDAL ACTIVITY IN MAIZE
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6320100artis Agribusiness Biotechnology Research,
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/547,422
; FILING DATE: 11-Apr-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/459,595
; FILING DATE: 02-JUN-1995
; APPLICATION NUMBER: US 07/951,715
; FILING DATE: 25-SEP-1992
; APPLICATION NUMBER: US 07/772,027
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: S-18805H
; TELECOMMUNICATION INFORMATION:

```



```

RESULT 39
US-07-951-715A-26
; Sequence 26, Application US/07951715A
; Patent No. 5625136
; GENERAL INFORMATION:
; APPLICANT: Koziel, Michael G.
; APPLICANT: Desai, Nalin M.
; APPLICANT: Lewis, Kelly S.
; APPLICANT: Kramer, Vance C.
; APPLICANT: Warren, Gregory W.
; APPLICANT: Evola, Stephen V.
; APPLICANT: Crossland, Lyle D.
; APPLICANT: Wright, Martha S.
; APPLICANT: Merlin, Ellis J.
; APPLICANT: Launis, Karen L.
; APPLICANT: Rothstein, Steven J.
; APPLICANT: Bowman, Cindy G.
; APPLICANT: Dawson, John L.
; APPLICANT: Dunder, Erik M.
; APPLICANT: Pace, Gary M.
; APPLICANT: Suttie, Janet L.
; TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
; TITLE OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: New York
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/951,715A
; FILING DATE: 25-SEP-1992
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/772,027
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Spruill, W. Murray
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: S-18805/A/CGC 1577/CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)541-8615
; TELEFAX: (919)541-8689
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4165 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1416..1425
; OTHER INFORMATION: /note= "start of mRNA"
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1481..2366
; FEATURE:
; NAME/KEY: intron
; LOCATION: 2367..2449
; FEATURE:
; NAME/KEY: exon
; LOCATION: 2450..2602
; FEATURE:
; NAME/KEY: intron

```

	Query Match	1.1%;	Score 35.6;	DB 4;	Length 4162;
	Best Local Similarity	54.6%;	Pred. No. 8.3;		
	Matches	71;	Conservative	0;	Mismatches 59; Indels 0; Gaps 0;
QY	38	AACCGGCTGCCATGGGGTTAGCTGTTCACAACACACTTTCTGCCTGCTTAGCGGGCGGCG	97		
Db	1581	AACGGGCGTGTCTCGTCGGCGGTGGCTGCTCCGCGTGTGTGCCGGTGGTGGTGGCG	1640		
QY	98	CGGGCACTTTCGCGCCGACTTCAATCGAGCGCGCACCGGTATCGCGACGAACAACGACAG	157		
Db	1641	CGGCGACGAGAACCGCGCTCACCCACCGCGCGCGCCAGGCGCAGTCCGGCAGCAAC	1700		
QY	158	CAACAACAGC	167		
Db	1701	CGGCGGGCGG	1710		



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 27, 2003, 16:39:33 ; Search time 434 Seconds  
(without alignments)  
16625.352 Million cell updates/sec

Title: US-09-830-433A-7  
Perfect score: 3204  
Sequence: 1 atggaacgaccccaacatt.....gcgtaggctacgggtctga 3204

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues 4370478  
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_101002:\*

- 1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*
- 2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*
- 3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*
- 4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*
- 5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*
- 6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*
- 7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*
- 8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*
- 9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*
- 10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*
- 11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*
- 12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*
- 13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*
- 14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*
- 15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*
- 16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*
- 17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*
- 18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*
- 19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*
- 20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*
- 21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*
- 22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*
- 23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*
- 24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	3204	100.0	3204	21	AAA15299		DNA encoding a pol
2	3062.2	95.6	13423	21	AAA81527		N. meningitidis pa
3	3062.2	95.6	349980	21	AAF21612		Neisseria meningit
4	2986.2	93.2	4218	22	AAS43905		Neisseria meningit
5	2986.2	93.2	4218	22	AAD17058		N. meningitidis st
6	2983.8	93.1	4335	22	AAS43899		Neisseria meningit
7	2983.8	93.1	4335	22	AAD17052		N. meningitidis st
8	2983.2	93.1	3939	22	AAS43880		Neisseria meningit
9	2983.2	93.1	3939	22	AAD17045		N. meningitidis st

10	2983.2	93.1	4170	22	AAS43902	Neisseria meningit
11	2983.2	93.1	4170	22	AAD17055	N. meningitidis st
12	2971.8	92.8	3939	22	AAS43875	Neisseria meningit
13	2971.8	92.8	3939	22	AAD17040	N. meningitidis st
14	2971.8	92.8	4179	22	AAS43877	Neisseria meningit
15	2971.8	92.8	4179	22	AAD17042	N. meningitidis st
16	2971.8	92.8	4344	22	AAS43876	Neisseria meningit
17	2971.8	92.8	4425	22	AAS43874	Neisseria meningit
18	2971.8	92.8	4425	22	AAD17039	N. meningitidis st
19	1946.6	60.8	3798	22	AAD17041	Neisseria meningit
20	1349	42.1	1365	21	AAS33301	Neisseria meningit
21	1342.6	41.9	1365	21	AAS33300	Neisseria meningit
22	1288.2	40.2	1365	21	AAS33299	Neisseria gonorrh
23	494.2	15.4	526	21	AAS33298	Neisseria meningit
24	470.6	14.7	528	21	AAS33297	Neisseria meningit
25	433.4	13.5	537	21	AAS33296	Neisseria gonorrh
26	270.8	8.5	286	19	AAV03603	Neisseria meningit
27	270.8	8.5	286	21	AAAL5396	Genomic fragment o
28	74.4	2.3	577	21	AA82003	N. meningitidis pa
29	65.4	2.0	476	21	AA81823	N. meningitidis pa
30	62	1.9	10732	21	AAAL0594	Gene encoding a su
31	49.6	1.5	913	18	AAAT91476	Mycobacterium tube
32	49.6	1.5	913	18	AAAT91413	Mycobacterium tube
33	49.6	1.5	913	19	AAV64462	M. tuberculosis im
34	49.6	1.5	913	19	AAV44354	Mycobacterium tube
35	49.6	1.5	913	20	AAZ19264	M. tuberculosis an
36	49.6	1.5	913	20	AAZ19052	M. tuberculosis re
37	49.6	1.5	4403765	22	AAI99683	Mycobacterium tube
38	49.6	1.5	4411529	22	AAI99682	Mycobacterium tube
39	48.4	1.5	750	24	ABK74700	Bacillus lichenifo
40	45.2	1.4	4403765	22	AAI99683	Mycobacterium tube
41	45.2	1.4	4411529	22	AAI99682	Mycobacterium tube
42	45	1.4	4384	22	AAF57288	M. bovis Dalton 2d
43	41.8	1.3	3946	18	AAAT93610	Mycobacterium tube
44	41.4	1.3	2949	21	AAD00024	Degenerate DNA enc
45	41.4	1.3	2949	24	AAD37455	Human Zfsta2 degen

ALIGNMENTS

RESULT 1  
AAA15299  
ID AAA15299 standard; DNA; 3204 BP.  
XX  
AC AAA15299;  
XX  
DT 04-SEP-2000 (first entry)  
XX  
DE DNA encoding a polypeptide of a Neisseria pathogenic strain.  
XX  
DE Pathogenic strain; Neisseria; vaccine; Neisseria infection; ss.  
XX  
OS Neisseria meningitidis.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..3204  
ET /\*tag= a  
XX  
XX WO200026375-A2.  
XX  
XX 11-MAY-2000.  
XX  
XX 28-OCT-1999; 99WO-FR02643.  
XX  
XX 30-OCT-1998; 98FR-0013693.  
XX  
XX (INMR ) PASTEUR MERIEUX SERUMS & VACCINS SA.  
XX  
XX (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.  
XX  
XX Aujaume L, Bouchardon A, Renauld-Mongenie G, Rokbi B, Nassif X;  
XX Tinsley C, Perrin A;

2200  
R/V

DR WPI: 2000-365622/31.  
XX P-PSDB: AAY93268.

PT New polypeptide specific for pathogenic Neisseria useful in therapeutic  
or preventative vaccines and for diagnosis

XX Claim 2; Page 61-65; 187pp; French.

CC The present sequence encodes a protein that is specific for pathogenic  
CC strains of Neisseria. The polynucleotides, polypeptides, or their  
CC antigenic fragments, are used in vaccines to treat or protect against  
CC Neisseria infections, particularly by N. meningitidis. The  
CC polynucleotide sequence is also used for recombinant production of  
CC the polypeptide and to produce attenuated Neisseria strains that  
CC overexpress it, or express it in a non-toxic mutant form.

XX Sequence 3204 BP; 831 A; 909 C; 899 G; 565 T; 0 other;

Query Match 100.0%; Score 3204; DB 21; Length 3204;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCGNACACCCACCTTCTCCCTACAAAACCTTTCAAAACCGGCTGCCATGGCGTTAGCT 60  
DB 1 ATGCGNACACCCACCTTCTCCCTACAAAACCTTTCAAAACCGGCTGCCATGGCGTTAGCT 60  
QY 61 GTTGCAACAACTTCTGCTGCTTAGCGGGGGGGGCGGCTTCTGCGCCCGACTTC 120  
DB 61 GTTGCAACAACTTCTGCTGCTTAGCGGGGGGGGCGGCTTCTGCGCCCGACTTC 120  
QY 121 AATGACGGCGGACCGGTTATCGGACGACAAACAGACAGCAACAGCAAGCAATCAGCAGCA 180  
DB 121 AATGACGGCGGACCGGTTATCGGACGACAAACAGACAGCAACAGCAAGCAATCAGCAGCA 180  
QY 181 GTATCTTACGCGGTATCAAGAACGAAATGTCAAGACAGAGCATGCTGTGCGCGT 240  
DB 181 GTATCTTACGCGGTATCAAGAACGAAATGTCAAGACAGAGCATGCTGTGCGCGT 240  
QY 241 CGGGATGACGTTGCGGTTACAGACAGGATGCCAAATCAATGCCCGCCCGGAACTCTG 300  
DB 241 CGGGATGACGTTGCGGTTACAGACAGGATGCCAAATCAATGCCCGCCCGGAACTCTG 300  
QY 301 CATACGGGAGACTTTACAAACCCCAATGACGCATACAAAGATTTGATCAACCTCAAACT 360  
DB 301 CATACGGGAGACTTTACAAACCCCAATGACGCATACAAAGATTTGATCAACCTCAAACT 360  
QY 361 GCAATTTGAAGCAGGCTATACAGACGCGGGTAGAGTAGGTATCGTCATACAGGCGAA 420  
DB 361 GCAATTTGAAGCAGGCTATACAGACGCGGGTAGAGTAGGTATCGTCATACAGGCGAA 420  
QY 421 TCCGTCGGCAGCATATCCTTTCCGGAACCTGTATGGCAGAAAAACACAGGCTATACAGAA 480  
DB 421 TCCGTCGGCAGCATATCCTTTCCGGAACCTGTATGGCAGAAAAACACAGGCTATACAGAA 480  
QY 481 AATTACAAAACCTATACGGCGTATATCGGAGAGAGCCCTGAGACAGGCGGCTAA 540  
DB 481 AATTACAAAACCTATACGGCGTATATCGGAGAGAGCCCTGAGACAGGCGGCTAA 540  
QY 541 GACATTTAAAGCTTCTTTCCAGCATGAGCGCTTATAGAGACTGAAGCAAGCCGAGCAT 600  
DB 541 GACATTTAAAGCTTCTTTCCAGCATGAGCGCTTATAGAGACTGAAGCAAGCCGAGCAT 600  
QY 601 ATCCGCGACGTAAGAAATCGSACACATCGATGTGGTCTCCCATATTTATGGCGGCGT 660  
DB 601 ATCCGCGACGTAAGAAATCGSACACATCGATGTGGTCTCCCATATTTATGGCGGCGT 660  
QY 661 TCCGTGGACGAGACCTCGAGCGGTATTTGGCCCGGATCGACGCTACACATAATGAAT 720  
DB 661 TCCGTGGACGAGACCTCGAGCGGTATTTGGCCCGGATCGACGCTACACATAATGAAT 720  
QY 721 ACGCATGTGGAAACCAAGAAATTAATGTCTGACGACCATCCGCAATGCATGGGTCAAG 780  
DB 721 ACGCATGTGGAAACCAAGAAATTAATGTCTGACGACCATCCGCAATGCATGGGTCAAG 780

QY 781 CTGGCGAACGTTGGCTGCGCATCTCAATAACAGTCTTTTGAACAACATCGAGGCGAGC 840  
DB 781 CTGGCGAACGTTGGCTGCGCATCTCAATAACAGTCTTTTGAACAACATCGAGGCGAGC 840  
QY 841 ACTGCCGACCAATTTCCAAATAGCCAAATTCGGAGGAGAGTACCGCCAAAGCGTTGCTGCC 900  
DB 841 ACTGCCGACCAATTTCCAAATAGCCAAATTCGGAGGAGAGTACCGCCAAAGCGTTGCTGCC 900  
QY 901 TATTCGCGGGTGATAAACAAGACGAGGATTCGCCCTGATGCAACAGAGGATTCACGC 960  
DB 901 TATTCGCGGGTGATAAACAAGACGAGGATTCGCCCTGATGCAACAGAGGATTCACGC 960  
QY 961 AACTTTGCTCTACCATCCGTAATAAACAACATGCTTTTTCATTTTTCGCGAAGCAATGAC 1020  
DB 961 AACTTTGCTCTACCATCCGTAATAAACAACATGCTTTTTCATTTTTCGCGAAGCAATGAC 1020  
QY 1021 GCACAAGCTCAGCCCAACACATGACCTATTTGCCATTTTATGAAAAAGATGCTCAAAA 1080  
DB 1021 GCACAAGCTCAGCCCAACACATGACCTATTTGCCATTTTATGAAAAAGATGCTCAAAA 1080  
QY 1081 GGCATTATACAGTCCGAGGCTAGACCGAGTGGAGAAAAGTTCAATGGCTTCCAAACAT 1140  
DB 1081 GGCATTATACAGTCCGAGGCTAGACCGAGTGGAGAAAAGTTCAATGGCTTCCAAACAT 1140  
QY 1141 TCGCGAATTTACTGCCATGCTGCTATCGGCACCTATGAAGCAAGGCTCGGTTTCACC 1200  
DB 1141 TCGCGAATTTACTGCCATGCTGCTATCGGCACCTATGAAGCAAGGCTCGGTTTCACC 1200  
QY 1201 CGTACAAAACCGGATTCAAATTCGCGGAAACATCTTTTCGACCCATCGTAACCGGACG 1260  
DB 1201 CGTACAAAACCGGATTCAAATTCGCGGAAACATCTTTTCGACCCATCGTAACCGGACG 1260  
QY 1261 GCGGCTCTGCTCGACAAATACCCGTTGATGAGCAACAGCAACCTGCTTACCAAGCTG 1320  
DB 1261 GCGGCTCTGCTCGACAAATACCCGTTGATGAGCAACAGCAACCTGCTTACCAAGCTG 1320  
QY 1321 CTGACAACGGCTCAGGACATCGGTGCTGCGGTGGACAGCAAGTTTCGGTGGGAGCTG 1380  
DB 1321 CTGACAACGGCTCAGGACATCGGTGCTGCGGTGGACAGCAAGTTTCGGTGGGAGCTG 1380  
QY 1381 CTGGATGCGGGTAAAGCCATGAACGACCGCTCTTTCGTTCCGGGACCTTTACCGC 1440  
DB 1381 CTGGATGCGGGTAAAGCCATGAACGACCGCTCTTTCGTTCCGGGACCTTTACCGC 1440  
QY 1441 GATACGAAAGGTACATCCGATATTCCTTCTCCGTAACGACATTTTCAGGCGAGG 1500  
DB 1441 GATACGAAAGGTACATCCGATATTCCTTCTCCGTAACGACATTTTCAGGCGAGG 1500  
QY 1501 GGCCTGATCAAAAAGGCGGACGCAACTGCAACTGCGGCAACAAACCTATTCAGGCG 1560  
DB 1501 GGCCTGATCAAAAAGGCGGACGCAACTGCAACTGCGGCAACAAACCTATTCAGGCG 1560  
QY 1561 AAAACCATTTATCGAAGCGGTTTCGCTGCTGTTGTAGGCAACAAACCTATTCAGGCG 1620  
DB 1561 AAAACCATTTATCGAAGCGGTTTCGCTGCTGTTGTAGGCAACAAACCTATTCAGGCG 1620  
QY 1621 GTGCAAAACAAAGGTGCGCTGATTTATAACGGGCGGATTCGCGGCTAGCTTGAACAG 1680  
DB 1621 GTGCAAAACAAAGGTGCGCTGATTTATAACGGGCGGATTCGCGGCTAGCTTGAACAG 1680  
QY 1681 GACGCGATTTCTATCTGCGAGATACCGACCGATCCGGCGCAACAAACCTGTCACATC 1740  
DB 1681 GACGCGATTTCTATCTGCGAGATACCGACCGATCCGGCGCAACAAACCTGTCACATC 1740  
QY 1741 AAAGCGGATCTCGAGTGGCGGCGGAGTACGCTGTACACAGCTTTGGGCAAACTGCTG 1800  
DB 1741 AAAGCGGATCTCGAGTGGCGGCGGAGTACGCTGTACACAGCTTTGGGCAAACTGCTG 1800  
QY 1801 AAAGTGGAGCTACGCGATGACCGGCGCAAGCTGTACATCTCGGCGACGCGGCAAGG 1860  
DB 1801 AAAGTGGAGCTACGCGATGACCGGCGCAAGCTGTACATCTCGGCGACGCGGCAAGG 1860

Db	2941	GATAAACCGTCTCTGTTTGCACGGCGGCGTGTGAACGGGACCTGAACGACGCGACTAC	3000
Qy	3001	ACGATAACGGCGCGCTTTACCGGCGGACTTCAGCAACCGCAACGCGGGGCACGCAAT	3060
Db	3001	ACGATAACGGCGCGCTTTACCGGCGGACTTCAGCAACCGCAACGCGGGGCACGCAAT	3060
Qy	3061	ATGCCGCACACCCGCGCTGTTGCCGTCTGGGCGGGGATGTGCAATTCCGCAACGGCTGG	3120
Db	3061	ATGCCGCACACCCGCGCTGTTGCCGTCTGGGCGGGGATGTGCAATTCCGCAACGGCTGG	3120
Qy	3121	AACGGCTTGGCAGCTTACAGCTACGCCGCTTCCAAACAGTACGGCAACACGAGGGACGA	3180
Db	3121	AACGGCTTGGCAGCTTACAGCTACGCCGCTTCCAAACAGTACGGCAACACGAGGGACGA	3180
Qy	3181	GTCGGCTAGGCTACCGGTTCTGA	3204
Db	3181	GTCGGCTAGGCTACCGGTTCTGA	3204

AAA81527/c  
ID AAA81527 standard; DNA: 13423 BP.

XX  
DT 04-DEC-2000 (first entry)

KW *Neisseria meningitidis*; *Neisseria gonorrhoeae*; genome; immunogenic;  
KW antigen; vaccine; diagnosis; infection; antibacterial; identification;  
KW *Meningococcus B*; MenB; ds.  
XX  
XX  
OS *Neisseria meningitidis*.

PD 20-APR-2000

XX  
XX

30-APR-1999; 9905-0132068.  
XX

PI Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;

DR WPI; 2000-318079/27.

PS Claim 7; Page 1439-1443; 1760pp; English.

The present invention describes methods of obtaining immunogenic proteins from *Neisseria* genomic sequences. AA811453 to AA82414 represent specifically claimed *Neisseria meningitidis* genomic DNA sequences; AA81260 to AA81303 and AA825620 to AA825663 represent *Neisseria* DNA sequences and their corresponding proteins; AA81254 to AA81259 and AA81304 to AA81321 represent PCR primers used in the isolation of *Neisseria meningitidis* DNA sequences; and AA81322 to AA81452 represent *Neisseria meningitidis* MenB polynucleotide ORF sequences, which are all used in the exemplification of the present invention. The nucleic acid sequences, protein sequences, and antibodies against them, can be used in the manufacture of a composition. The composition can be used as a medicament (or in the manufacture of a medicament) for treating, preventing or diagnosing infection due to *Neisserial* bacteria. For example, some of the identified proteins could

be components of vaccines against *Meningococcus B*; against all serotypes; and/or against all pathogenic *Neisseriae*. Identification of sequences from the bacterium will also facilitate production of biological probes, particularly organism-specific probes. Attempts to make efficacious *Meningococcus B* vaccines have failed mainly due to antigen tolerance. Multivalent vaccines have also been tried but none have successfully overcome antigenic variability. The provision of further, complete sequences may provide an opportunity to identify secreted or surface exposed proteins that may be presumed targets for the immune system and which are not antigenically variable or at least more conserved than other more variable regions.

Sequence 13423 BP; 2916 A; 3336 C; 3842 G; 3329 T; 0 other;

Query Match	95.6%;	Score	3062.2;	DB	21;	Length	13423;
Best Local Similarity	97.1%;	Pred.	No. 0;				
Matches	3156;	Conservative	0;	Mismatches	48;	Indels	45;
Gaps	2;						
QY	1	ATGCGAAGACCCCAACCTTCCTACAAAACCTTTCAAAACCGGCTGCCATGCGGTAGCT	60				
Db	10889	ATGCGAAGACCCCAACCTTCCTACAAAACCTTTCAAAACCGGCTGCCATGCGGTAGCT	10830				
QY	61	GTTCGAACAACACTTTCCTGCTTAGCGCGGC---GGCGGCACCTCTCGCGCCGAC	117				
Db	10829	GTTCGAACAACACTTTCCTGCTTAGCGCGGC---GGCGGCACCTCTCGCGCCGAC	10770				
QY	118	TTCAATGACGCGGCACCGGTATCGGACCAACAGACAGCAACAACAGCAAAATCAGCA	177				
Db	10769	TTCAATGACGCGGTACCGGTATCGGACCAACAGACAGCAACAACAGCAAAATCAGCA	10710				
QY	178	GCAGTATCTTACGCGGTATCAAGAACCAATGTGCAAGACAGAAAGTCTGTGTC	237				
Db	10709	GCAGTATCTTACGCGGTATCAAGAACCAATGTGCAAGACAGAAAGTCTGTGTC	10650				
QY	238	GTCTGGATGACGTTCGGTTACAGACGGGATGCCAAATCAATGCCCGCCCGGAT	297				
Db	10649	GTCTGGATGACGTTCGGTTACAGACGGGATGCCAAATCAATGCCCGCCCGGAT	10590				
QY	298	CTGCATACCGGAGACTTTACAAACCAATGACGATACAAAGAAATTTGATCAACCTCAA	357				
Db	10589	CTGCATACCGGAGACTTTCCAAACCAATGACGATACAAAGAAATTTGATCAACCTCAA	10530				
QY	358	CTGCAATTTGAAGCAGGTATACAGACGCGGGTAGGTTAGGTATCGTCAATACAGC	417				
Db	10529	CTGCAATTTGAAGCAGGTATACAGACGCGGGTAGGTTAGGTATCGTCAATACAGC	10470				
QY	418	GAATCCGTCGGCAGCATATCTTTCCGAACCTGTATGCAGAAAGAAACACGCTATAC	477				
Db	10469	GAATCCGTCGGCAGCATATCTTTCCGAACCTGTATGCAGAAAGAAACACGCTATAC	10410				
QY	478	GAATAATCAAAAACCTATACGGCGTATATCGGAAGAGCGCCTGAAGACGGAGCGGT	537				
Db	10409	GAATAATCAAAAACCTATACGGCGTATATCGGAAGAGCGCCTGAAGACGGAGCGGT	10350				
QY	538	AAAGACATTAAGCTTCTTTCAGCATGAGCGCGTTATAGACACTGAAGCAAGCCGAGC	597				
Db	10349	AAAGACATTAAGCTTCTTTCAGCATGAGCGCGTTATAGACACTGAAGCAAGCCGAGC	10290				
QY	598	GATATCCGCGCAGTAAAGAAATCGACATCGATGTGCTCCCATATTTATGGCGGG	657				
Db	10289	GATATCCGCGCAGTAAAGAAATCGACATCGATGTGCTCCCATATTTATGGCGGG	10230				
QY	658	CGTCCGTCGGACGACCTCGAGCGGTATTTGCGCCCGATGCGACGCTACACATAATG	717				
Db	10229	CGTCCGTCGGACGACCTCGAGCGGTATTTGCGCCCGATGCGACGCTACACATAATG	10170				
QY	718	AATAGCATATGGAACCAAGCAAAATATGTCTGAGCCCATCCGCAATGCATGGGTC	777				
Db	10169	AATAGCATATGGAACCAAGCAAAATATGTGTGAGCCCATCCGCAATGCATGGGTC	10110				
QY	778	AAGCTTGGCGAAGCTGGCGTGGCATCGTCAATAACAGTTTTTGAACAACATCGAGGCA	837				
Db	10109	AAGCTTGGCGAAGCTGGCGTGGCATCGTCAATAACAGTTTTTGAACAACATCGAGGCA	10050				

QY	838	GGCACTGCGGACCACTTTCCAAATAGCCAAATTCGGAGGACGAGTACCGCAACGCTTGCTC	897				
Db	10049	GGCACTGCGGACCTTTTCCAAATAGCCAAATTCGGAGGACGAGTACCGCAACGCTTGCTC	9990				
QY	898	GCCTATTCGGCGGTGATAAACAAGACAGAGGTATCCGCGCTGATCAACAGAGCGATAC	957				
Db	9989	GACTATTCCGGCGGTGATAAACAAGACAGAGGTATCCGCGCTGATCAACAGAGCGATAC	9930				
QY	958	GGCAACTTGTCTACCAATCCGTAATAAACAACATGCTTTTCATTTTTCGCAAGCAAT	1017				
Db	9929	GGCAACTTGTCTACCAATCCGTAATAAACAACATGCTTTTCATTTTTCGCAAGCAAT	9870				
QY	1018	GACGACAAAGCTGACGCCCAACACACTGACCTTATTTGCCATTTTATGAAAAGATGCTCAA	1077				
Db	9869	GACGACAAAGCTGACGCCCAACACATATGCCCTATTTGCCATTTTATGAAAAGATGCTCAA	9810				
QY	1078	AAAGCATATACAGTCGAGGCTAGACCGCATAGACGAGTGGAGAAAATTC	1126				
Db	9809	AAAGCATATACAGTCGAGGCTAGACCGCATAGACGAGTGGAGAAAATTC	9750				
QY	1127	-----ATGGCTCCAACCATTTGGGAATTTACTGCC	1155				
Db	9749	TATGAGAAACCGGGTACAGAACCGCTTGAGTATGGCTCCAACCATTTGGGAATTTACTGCC	9690				
QY	1156	ATGTGCTGCTATCGGCACCTTATGAAGCAAGCTCCGTTTCAACCGGTACAAACCCGAT	1215				
Db	9689	ATGTGCTGCTATCGGCACCTTATGAAGCAAGCTCCGTTTCAACCGGTACAAACCCGAT	9630				
QY	1216	CAATTTGCGGAAACATCTTTTCCGACCCATCTGTAACCGGACGCGGCTCTGCTGTG	1275				
Db	9629	CAATTTGCGGAAACATCTTTTCCGACCCATCTGTAACCGGACGCGGCTCTGCTGTG	9570				
QY	1276	CAGAAATACCGTGTGATGAGCAACGACACCTGCTTACACGCTCTGACACGCTCAG	1335				
Db	9569	CAGAAATACCGTGTGATGAGCAACGACACCTGCTTACACGCTCTGACACGCTCAG	9510				
QY	1336	GACATTCGCTGAGTGGGTGGACAGCAAGTTTCGGCTGGGACTCTGATCGGGTAA	1395				
Db	9509	GACATTCGCTGAGTGGGTGGACAGCAAGTTTCGGCTGGGACTCTGATCGGGTAA	9450				
QY	1396	GCCATGAACGACCGCGCTCTTCCGTTCCGCGACTTTACGCGCGATACGAAAGTACA	1455				
Db	9449	GCCATGAACGACCGCGCTCTTCCGTTCCGCGACTTTACGCGCGATACGAAAGTACA	9390				
QY	1456	TCCGATTTGCTACTCTCTTCCGTAACGACATTTTCAGGACGCGGCTGATCAAAAA	1515				
Db	9389	TCCGATTTGCTACTCTCTTCCGTAACGACATTTTCAGGACGCGGCTGATCAAAAA	9330				
QY	1516	GGCGGACCAACTGCAACTGCACGGCAACACACTATACGGGCAAAACCAATTTATCGAA	1575				
Db	9329	GGCGGACCAACTGCAACTGCACGGCAACACACTATACGGGCAAAACCAATTTATCGAA	9270				
QY	1576	GGCGTTTCGCTGTTGTAGCGCAACAAATCGGATATGCGGCTGAAACCAAGT	1635				
Db	9269	GGCGTTTCGCTGTTGTAGCGCAACAAATCGGATATGCGGCTGAAACCAAGT	9210				
QY	1636	GGCGTTTATTAACGGGGGCGCATCCGCGGTAGCCTGAACAGGACGGCATTTCTAT	1695				
Db	9209	GGCGTTTATTAACGGGGGCGCATCCGCGGTAGCCTGAACAGGACGGCATTTCTAT	9150				
QY	1696	CTGGCAGATACCGACCGATCCGGCGCAACACCGGTGACATCAAAAGGCGATCTGAG	1755				
Db	9149	CTGGCAGATACCGACCAATCCGCGCAACACCGGTACACATCAAAAGGCGATCTGAG	9090				
QY	1756	CTGGCGCGGAGGTACCGCTGTACACAGCTTTGGGCAAACTGCTGAAAGTGACGGTAC	1815				
Db	9089	CTGGCGCGGAGGTACCGCTGTACACAGCTTTGGGCAAACTGCTGAAAGTGACGGTAC	9030				
QY	1816	GCGATGACCGCGGCAAGCTGTACATGTCCGACCGCGCAAGGGGCGGCTATCTCAAC	1875				
Db	9029	GCGATGACCGCGGCAAGCTGTACATGTCCGACCGCGCAAGGGGCGGCTATCTCAAC	8970				

QY 1876 CGTACCGGACAAAGTGTTCCTTCCTTCAGTGGCGCCAAATCGGGCGGATATTCTTTC 1935  
Db 1877 TTTTACCGGCGGCGCTTTCCTTCCTTCAGTGGCGCCAAATCGGGCGGATATTCTTTC 8910  
QY 1936 TTTACCAAAATCATGAAACCGGAGTGTCTGCTGGCTTCCTTCGACAGCGTCGAAACAA 1995  
Db 1937 TTTACCAAAATCATGAAACCGGAGTGTCTGCTGGCTTCCTTCGACAGCGTCGAAACAA 8850  
QY 1996 GCGGGCAGTGAAGCGGACACAGCTGTCTTATATGTCTCGCGGCAATCGGGCAGCGACT 2055  
Db 1997 GCGGGCAGTGAAGCGGACACAGCTGTCTTATATGTCTCGCGGCAATCGGGCAGCGACT 8790  
QY 2056 GTTTCGGCAGCGCACAATTCGCGGCCCGCGGTCTGAAACACGCGCTAGAACAGGCGCG 2115  
Db 2057 GTTTCGGCAGCGCACAATTCGCGGCCCGCGGTCTGAAACACGCGCTAGAACAGGCGCG 8730  
QY 2116 AGCAATCTGGAAACCTGATGTCGAACCTGGATGCTCTCGAATCATCCGCAACCCGAG 2175  
Db 2117 AGCAATCTGGAAACCTGATGTCGAACCTGGATGCTCTCGAATCATCCGCAACCCGAG 8670  
QY 2176 ACGGTTGAAACTGCGCGCCGCGACCGCACAGATATCGCGGCATCGCCCTACGCGCA 2235  
Db 2177 ACGGTTGAAACTGCGCGCCGCGACCGCACAGATATCGCGGCATCGCCCTACGCGCA 8610  
QY 2236 ACTTTCGCGCAGCGCAGCGCTACAGCATGCGAATGCGCGCAGCGGTGTACGCATCTTC 2295  
Db 2237 ACTTTCGCGCAGCGCAGCGCTACAGCATGCGAATGCGCGCAGCGGTGTACGCATCTTC 8550  
QY 2296 ACAGTCTCGCGCTACCGCTATGCGCGACAGTACCGCGCCCGCATCGCATATGCAAGGA 2355  
Db 2297 ACAGTCTCGCGCTACCGCTATGCGCGACAGTACCGCGCCCGCATCGCATATGCAAGGA 8490  
QY 2356 CGCGGCTGAAAGCCGCTATCGGACGCGGTGGACCAACAGCTACGCGTCTGCGCGTCATC 2415  
Db 2357 CGCGGCTGAAAGCCGCTATCGGACGCGGTGGACCAACAGCTACGCGTCTGCGCGTCATC 8430  
QY 2416 GCGCAAAACCAACAGGACGCTGGAACGCTGGGAACAGGGCGGTGTTGAAGCAAAATGCGC 2475  
Db 2417 GCGCAAAACCAACAGGACGCTGGAACGCTGGGAACAGGGCGGTGTTGAAGCAAAATGCGC 8370  
QY 2476 GGCAGTACCAAAACCGTGGCATTCGCGGAAACCGGGGAAATACGACAGCGCGCC 2535  
Db 2477 GGCAGTACCAAAACCGTGGCATTCGCGGAAACCGGGGAAATACGACAGCGCGCC 8310  
QY 2536 ACACGCGGATGGGACACACACATGACGAGCGAAACAGTGCAAATGCAAAACCGACAGC 2595  
Db 2537 ACACGCGGATGGGACACACACATGACGAGCGAAACAGTGCAAATGCAAAACCGACAGC 8250  
QY 2596 ATTAGTCTGTTTCAGGCATACGCGACGATGCGGGCGATATCGGCTATCTCAAAGCCCTG 2655  
Db 2597 ATTAGTCTGTTTCAGGCATACGCGACGATGCGGGCGATATCGGCTATCTCAAAGCCCTG 8190  
QY 2656 TTCTCTACGGAGCTTACAAAACACATCAGCCGCGAGCAGCGGTGCGGACGAATGCG 2715  
Db 2657 TTCTCTACGGAGCTTACAAAACACATCAGCCGCGAGCAGCGGTGCGGACGAATGCG 8130  
QY 2716 GAAGGACGCTCAACGCGACGCTGATGCGGCGCACTGGGCGGTGTCAACGTTCCG 2775  
Db 2717 GAAGGACGCTCAACGCGACGCTGATGCGGCGCACTGGGCGGTGTCAACGTTCCG 8070  
QY 2776 TTTTCGCGCAACGGGAGATTGACGGTGAAGCGGTCTGCGCTACGACCTGTCTCAAACAG 2835  
Db 2777 TTTTCGCGCAACGGGAGATTGACGGTGAAGCGGTCTGCGCTACGACCTGTCTCAAACAG 8010  
QY 2836 GTTGCAATTCGCGGAAAGGACGCTGTTTGGGTGGAGCGGCAACGCTCATCTGAAGGC 2895  
Db 2837 GTTGCAATTCGCGGAAAGGACGCTGTTTGGGTGGAGCGGCAACGCTCATCTGAAGGC 7950  
QY 2896 ACACGCTGCGGACTCGCGGTCTGAAGCTCTGCAACCCCTTGAGCGCATTAAGCGCTCTG 2955  
Db 2897 ACCTGCTGCGACTCGCGGTCTGAAGCTCTGCAACCCCTTGAGCGCATTAAGCGCTCTG 7890  
QY 2956 TTTTGAACGCGCGGCGTGGAAACGCGACCTGAACGCGGCGGACTTACACGGTAAACGCGCGC 3015

Db 7889 TTTTGAACGCGCGGCGTGGAAACGCGACTGAACGCGGCGACTTACACGGTAAACGCGCGC 7830  
QY 3016 TTTTACCGCGCGGCTGACGCAACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3075  
Db 7829 TTTTACCGCGCGGCTGACGCAACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 7770  
QY 3076 CTGTTTCCGCGGCTGCGGCGGCGGATGTCGAATTCGGCAACCGCTTGGACGCGT 3135  
Db 7769 CTGTTTCCGCGGCTGCGGCGGCGGATGTCGAATTCGGCAACCGCTTGGACGCGT 7710  
QY 3136 TACAGTACGCGCGGTTCCAAACAGTACGCGCAACGCGGAGTTCGCGGTAGGCTAC 3195  
Db 7709 TACAGTACGCGCGGTTCCAAACAGTACGCGCAACGCGGAGTTCGCGGTAGGCTAC 7650  
QY 3196 CGGTTCTGA 3204  
Db 7649 CGGTTCTGA 7641  
RESULT 3  
AAF21612  
ID AAF21612 standard; DNA; 349980 BP.  
XX AAF21612;  
XX AC  
DT 13-MAR-2001 (first entry)  
XX  
DE Neisseria meningitidis B nucleotide sequence SEQ ID NO:113.  
XX  
KW Neisseria meningitidis; Neisseria gonorrhoeae; immunogenic; vaccine;  
KW diagnosis; antigen; detection; infection; gene therapy; antibacterial;  
ds.  
XX  
OS Neisseria meningitidis.  
XX  
PN WO200066791-A1.  
XX  
PD 09-NOV-2000.  
XX  
PF 08-MAR-2000; 2000WO-US05928.  
XX  
PR 30-APR-1999; 99US-0132068.  
PR 08-OCT-1999; 99WO-US23573.  
PR 28-FEB-2000; 2000GB-0004695.  
XX  
PA (CHIR ) CHIRON CORP.  
PA (GENO-) INST GENOMIC RES.  
XX  
PI Pizza M, Hickey E, Peterson J, Tettelin H, Venter JC, Maignani V;  
PI Galeotti C, Mora M, Ratti G, Scarselli M, Scariato V, Rappuoli R;  
PI Frazer CM, Grandi G;  
XX  
DR WPI; 2000-647603/62.  
XX  
PT Neisseria meningitidis B full length genome sequence and open reading  
PT frames are used to detect, treat and prevent Neisserial infections -  
XX  
PS Claim 7; Appendix A; 692pp; English.  
XX  
CC The present invention describes the full length genome of  
CC Neisseria meningitidis B (NMB). The sequences in AAF21544 and AAF21607  
CC to AAF21613 represent fragments of the NMB genomic sequence, as the  
CC sequence was too long to go in a record on its own it was split into 8  
CC sequences which overlap each other at the beginning and end of each  
CC sequence by 49980 bp (i.e. the last 49980 bp of AAF21544 is repeated at  
CC the beginning of AAF21607, the last 49980 bp of AAF21607 are repeated at  
CC the beginning of AAF21608, and so on). AAF21545 to AAF21588 encode the  
CC Neisseria proteins given in AAB58550 to AAB58593, and AAF21589 to  
CC AAF21606 represent PCR primers which are used in the exemplification of  
CC the present invention. The NMB genome and fragments from it have  
CC antibacterial activity, and can be used in vaccines and gene therapy.  
CC Neisseria nucleic acids, proteins and/or antibodies which binds to the

CC proteins can be used in compositions for treating or preventing infection  
CC due to Neisserial bacteria or as a diagnostic reagent for detecting the  
CC presence of Neisserial bacteria or of antibodies raised to Neisserial  
CC bacteria. Computers, computer memory, computer storage medium or computer  
CC databases can be used in a search to identify open reading frames (ORFs)  
CC or coding sequences within the NMB genome. The DNA sequences provide  
CC further opportunities to find antigenic or immunogenic proteins which are  
CC more effective in vaccines than the outer membrane proteins currently  
CC used.

XX Sequence 349980 BP; 86473 A; 95646 C; 85908 G; 81953 T; 0 other;

Query Match 95.6%; Score 3062.2; DB 21; Length 349980;  
Best Local Similarity 97.1%; Pred. No. 0;  
Matches 3156; Conservative 0; Mismatches 48; Indels 45; Gaps 2;

Qy 1 ATGCGAAGCGCCCAACCTTCCCTACAAAACTTCAAAACCGGCTGCCATGGCGTTAGCT 60  
Db 261412 ATGCGAAGCGCCCAACCTTCCCTACAAAACTTCAAAACCGGCTGCCATGGCGTTAGCT 261471  
Qy 61 GTTGCAACAACACTTTCGCTGCTTAGCGCGCGC---GGCGGCACCTCTCGCGCCGAC 117  
Db 261472 GTTGCAACAACACTTTCGCTGCTTAGCGCGCGCGGAGCGGCACCTCTCGCGCCGAC 261531  
Qy 118 TTCATGCGAGCGCGGACCGGTATCGGCAGCAACAGCAGAGCAACACGCGAAATTCAGCA 177  
Db 261532 TTCATGCGAGCGGTFACCGGTATCGGCAGCAACAGCAGAGCAACACGCGAAATTCAGCA 261591  
Qy 178 GCAGTATCTTAGCGCGGTATCAAGAACCAATGTGCAAGACAGAGCATGCTCTGTGCC 237  
Db 261592 GCAGTATCTTAGCGCGGTATCAAGAACCAATGTGCAAGACAGAGCATGCTCTGTGCC 261651  
Qy 238 GGTGCGGATGACGTTGCGGTATACAGACAGGATGCCAAATCAATGCCCGCCCGCCGAAT 297  
Db 261652 GGTGCGGATGACGTTGCGGTATACAGACAGGATGCCAAATCAATGCCCGCCCGCCGAAT 261711  
Qy 298 CTGCATACCGGAGACTTTACAAACCCAAATGACGCATACAAGAAATTTGATCAACCTCAA 357  
Db 261712 CTGCATACCGGAGACTTTTCCAAACCCAAATGACGCATACAAGAAATTTGATCAACCTCAA 261771  
Qy 358 CTTGCAATTAAGCAGGCTATACAGACGCGGGTAGAGGTAGGTATCGTCATACAGCG 417  
Db 261772 CTTGCAATTAAGCAGGCTATACAGACGCGGGTAGAGGTAGGTATCGTCACACAGCG 261831  
Qy 418 GAATCCGTCGGCAGCATATCCTTCCGCAACTGTATGCGAGAAAAAGACACGGCTATAAC 477  
Db 261832 GAATCCGTCGGCAGCATATCCTTCCGCAACTGTATGCGAGAAAAAGACACGGCTATAAC 261891  
Qy 478 GAAATTAACAAAACCTATACGGCGTATATGCGGAAGGAGCGCCTGAGACGGAGCGGT 537  
Db 261892 GAAATTAACAAAACCTATACGGCGTATATGCGGAAGGAGCGCCTGAGACGGAGCGGT 261951  
Qy 538 AAAGACATTAAGCTTCTTCCACGATGAGCGCGTTATAGACACTGAAGCAAGCGGAG 597  
Db 261952 AAAGACATTAAGCTTCTTCCACGATGAGCGCGTTATAGACACTGAAGCAAGCGGAG 262011  
Qy 598 GATATCCGCGACGTAAGAGAAATCGACACATCGATGTGGTCTCCCATATATTGGCGGG 657  
Db 262012 GATATCCGCGACGTAAGAGAAATCGACACATCGATTTGGTCTCCCATATATTGGCGGG 262071  
Qy 658 CGTTCGTTGGAGCGGACGCTCGAGCGGTATTCGCGCGGATGCGACGCTACACATAATG 717  
Db 262072 CGTTCGTTGGAGCGGACGCTCGAGCGGTATTCGCGCGGATGCGACGCTACACATAATG 262131  
Qy 718 ATACGATGATGGAAACCAAGAACCAATATGCTCGAGCGCATCGCAATCGATGGGTGTC 777  
Db 262132 ATACGATGATGGAAACCAAGAACCAATATGCTCGAGCGCATCGCAATCGATGGGTGTC 262191  
Qy 778 AAGCTTGGGCGAACGTTGGCGTGCCTCGTCAATAACAGTTTGGAAACACATCGAGGGCA 837  
Db 262192 AAGCTTGGGCGAACGTTGGCGTGCCTCGTCAATAACAGTTTGGAAACACATCGAGGGCA 262251  
Qy 838 GGCACGTGGCGACCATTTTCCAAATAGCCAAATTCGGAGGAGCAGTACCGGCAACGCTTGCTC 897

Db 262252 GGCACGTGGCGACCATTTTCCAAATAGCCAAATTCGGAGGAGCAGTAGTACCGCAACGCTTGCTC 262311  
Qy 898 GCCTATTCCGCGCGTGTATAAACAGAGAGGGGTATCGCCCTGATGCAACACAGCGATTAC 957  
Db 262312 GACATTCCGCGCGTGTATAAACAGAGAGGGGTATCGCCCTGATGCAACACAGCGATTAC 262371  
Qy 958 GGCACATTGTCTTACCACATCCGTAATAAAAAACATGCTTTTTCATTTCGCGCAAGCAAT 1017  
Db 262372 GGCACATTGTCTTACCACATCCGTAATAAAAAACATGCTTTTTCATTTCGCGCAAGCAAT 262431  
Qy 1018 GACGCAACAGCTCAGCCCAACACACACTGACCTTATGCAATTTTATGAAAAAGATGCTCAA 1077  
Db 262432 GACGCAACAGCTCAGCCCAACACACATATGCCCTATTTGCCATTTTATGAAAAAGAGCTCAA 262491  
Qy 1078 AAAGCATTTATCACAGTCGCGAGCGGTAGACCGCAGTGGAGAAAAGTTCAACGGGAAATG 1126  
Db 262492 AAAGCATTTATCACAGTCGCGAGCGGTAGACCGCAGTGGAGAAAAGTTCAACGGGAAATG 262551  
Qy 1127 -----ATGSGTCCAAACCATTTGCGGAATTTACTGCC 1155  
Db 262552 TATGAGAAACCGGTTACAGAACCGCTTGAGTATGGCTCCAACCATTTGCGGAATTTACTGCC 262611  
Qy 1156 ATGTGTCCTTATCGGCACCTATGAAGCAAGCGTCCGTTTTCACCCGTACAAACCCGATT 1215  
Db 262612 ATGTGTCCTTATCGGCACCTATGAAGCAAGCGTCCGTTTTCACCCGTACAAACCCGATT 262671  
Qy 1216 CAATTTGCGGAACATCCTTTTCGSCACCCATCGTAAACGGCACGCGGCTCTGCTGCTG 1275  
Db 262672 CAATTTGCGGAACATCCTTTTCGSCACCCATCGTAAACGGCACGCGGCTCTGCTGCTG 262731  
Qy 1276 CAGAAATCCCGTGGATGAGCAACGCAACCTGCGTACCACGCTGCTGACACGCGCTCAG 1335  
Db 262732 CAGAAATACCCGTGGATGAGCAACGCAACCTGCGTACCACGCTGCTGACACGCGCTCAG 262791  
Qy 1336 GACATCGTGCAGTCGGGTGGACAGCAAGTTTCGCTGGGACTGCTGGATCGCGGTAG 1395  
Db 262792 GACATCGTGCAGTCGGGTGGACAGCAAGTTTCGCTGGGACTGCTGGATCGCGGTAG 262851  
Qy 1396 GCCATGAACGGAACCGCGCTCTTTCGTCGCGACTTTACCGCGGATACGAAAGGTACA 1455  
Db 262852 GCCATGAACGGAACCGCGCTCTTTCGTCGCGACTTTACCGCGGATACGAAAGGTACA 262911  
Qy 1456 TCCGATATTGCTACTCTCCGTAAACGACATTTTCAGGACGGGCGGCTGATCAAAAA 1515  
Db 262912 TCCGATATTGCTACTCTCCGTAAACGACATTTTCAGGACGGGCGGCTGATCAAAAA 262971  
Qy 1516 GCGCGACGCAACTGCAACTGCACGGCAACACACTATACGGGCAAAACCATTTATCGAA 1575  
Db 262972 GCGCGACGCAACTGCAACTGCACGGCAACACACTATACGGGCAAAACCATTTATCGAA 263031  
Qy 1576 GCGGTTTCGCTGGTGTGTACGGCAACCAAAATCGGATATCGCGCTCGAAACCAAGGT 1635  
Db 263032 GCGGTTTCGCTGGTGTGTACGGCAACCAAAATCGGATATCGCGCTCGAAACCAAGGT 263091  
Qy 1636 GCGCTGATTTATAACGGGGCGGCTATCCGGCGGTAGCTGTAACAGCGACGGCATTTCTAT 1695  
Db 263092 GCGCTGATTTATAACGGGGCGGCTATCCGGCGGACGCTGAACAGCGACGGCATTTCTAT 263151  
Qy 1696 CTGGCAGATACCGACCGGATCCGGCGCAACGCAACCGTGCACATCAAGGGCATCTGCAG 1755  
Db 263152 CTGGCAGATACCGACCGGATCCGGCGCAACGCAACCGTGCACATCAAGGGCATCTGCAG 263211  
Qy 1756 CTGGCGGCGGAGGTACGCTGTACAGCTTTGGGCAAACTGCTGAAAAGTGACGCGTAG 1815  
Db 263212 CTGGCGGCGGAGGTACGCTGTACACAGCTTTGGGCAAACTGCTGAAAAGTGACGCGTAG 263271  
Qy 1816 GCGATGACCGCGGCAAGCTGTACATGTGCGCACGCGGCAAGGGCGAGGCTATCTCAAC 1875  
Db 263272 GCGATGACCGCGGCAAGCTGTACATGTGCGCACGCGGCAAGGGCGAGGCTATCTCAAC 263331  
Qy 1876 CGTACCGCAACAGCTGTTCCTTCTGAGTCCGCCCAAAATCGCGGCGGATTTCTTTC 1935



Db 263332 AGTACCGGACGAGTGTTCCTTCCTGAGTGGCGGCAAAATCGGGAGGATTATTCTTTC 263391  
QY 1936 TTCACAAACATCGAAACCGACGGTGTCTGCTGGCTTCCTCGACAGCGTCGAAAAACA 1995  
Db 263392 TTCACAAACATCGAAACCGACGGCGCTGCTGGCTTCCTCGACAGCGTCGAAAAACA 263451  
QY 1996 GGGGCGAGTGAAGGCGACAGCTGTCTCTATATGTCCGTGCGGGCAATCGGCGACGAGCT 2055  
Db 263452 GCGGGCAGTGAAGGCGACAGCTGTCTCTATATGTCCGTGCGGGCAATCGGCGACGAGCT 263511  
QY 2056 GCTTCGGCAGCGCAGATTCGCGGCGCGCGCTGCTGAACACGCCCTAGAACAGGCGCG 2115  
Db 263512 GCTTCGGCAGCGCAGATTCGCGGCGCGCGCTGCTGAACACGCCCTAGAACAGGCGCG 263571  
QY 2116 AGCAATCTGGAACACCTGATGTCGAACCTGGATGCTCCGGAATCATCCGCAACACCCGAG 2175  
Db 263572 AGCAATCTGGAACACCTGATGTCGAACCTGGATGCTCCGGAATCATCCGCAACACCCGAG 263631  
QY 2176 ACGGTTGAAACTCGGCGCGCGCCGACCGCACAGATATCCGGGCATCCGCCCTACGGCGCA 2235  
Db 263632 ACGGTTGAAACTCGGCGCGCGCCGACCGCACAGATATCCGGGCATCCGCCCTACGGCGCA 263691  
QY 2236 ACTTTCGGCGCAGCGCAGCGCTACAGCATGCGAATGCGCGCGACGGTGTACGCATCTTC 2295  
Db 263692 ACTTTCGGCGCAGCGCAGCGCTACAGCATGCGAATGCGCGCGACGGTGTACGCATCTTC 263751  
QY 2296 AACAGTCTCGCGCTACCGTCTATGCGGACAGTACCGCGCGCCCATCCGATATGCAAGGA 2355  
Db 263752 AACAGTCTCGCGCTACCGTCTATGCGGACAGTACCGCGCGCCCATCCGATATGCAAGGA 263811  
QY 2356 CCGCGGCTGAAAGCGCTATCGGACGGGTTGGACCAACAGCTACGGGTCTGCGCGTCATC 2415  
Db 263812 CCGCGGCTGAAAGCGCTATCGGACGGGTTGGACCAACAGCTACGGGTCTGCGCGTCATC 263871  
QY 2416 GCGCAACCCACAGACAGCGTGAAGCTGGGAACAGGGCGGTGTTGAAGCAAAATGCGC 2475  
Db 263872 GCGCAACCCACAGACAGCGTGAAGCTGGGAACAGGGCGGTGTTGAAGCAAAATGCGC 263931  
QY 2476 GCAGTACCCAAACCGTCGCGATTCGCGGAAACCGGAAATACGACAGCGCGC 2535  
Db 263932 GCAGTACCCAAACCGTCGCGATTCGCGGAAACCGGAAATACGACAGCGCGC 263991  
QY 2536 ACACCTGGGATGGACACAGCACATGAGCGAAACAGTGCAAATGCAAAACCGACAGC 2595  
Db 263992 ACACCTGGGATGGACACAGCACATGAGCGAAACAGTGCAAATGCAAAACCGACAGC 264051  
QY 2596 ATTAGTCTGTTGACGGCATACGGCAGATGCGGGCGATATCGGCTATCTCAAGGCCCTG 2655  
Db 264052 ATTAGTCTGTTGACGGCATACGGCAGATGCGGGCGATATCGGCTATCTCAAGGCCCTG 264111  
QY 2656 TTCTCCTACGGAGCTTACAAAACAGCATCAGCCGACGACCGGTGCGGACGAACATGG 2715  
Db 264112 TTCTCCTACGGAGCTTACAAAACAGCATCAGCCGACGACCGGTGCGGACGAACATGG 264171  
QY 2716 GAAGCAGCGTCAACGGCAGCTGTAGTSCAGTGGCGCAGCTGGGCGGTGTCAACGTTCCG 2775  
Db 264172 GAAGCAGCGTCAACGGCAGCTGTAGTSCAGTGGCGCAGCTGGGCGGTGTCAACGTTCCG 264231  
QY 2776 TTTGCGGCAACGGGATTTGACGGTGAAGCGGTGCTGGCTACGACCTGTCTCAACAG 2835  
Db 264232 TTTGCGGCAACGGGATTTGACGGTGAAGCGGTGCTGGCTACGACCTGTCTCAACAG 264291  
QY 2836 GATGCATTCGCGGAAAGGAGTGTGTTGGCTGAGCGGCAACAGCTCACTGAAGC 2895  
Db 264292 GATGCATTCGCGGAAAGGAGTGTGTTGGCTGAGCGGCAACAGCTCACTGAAGC 264351  
QY 2896 ACACCTGGTGGGCTCGGGGTCTGAAGCTGTGCAACCCCTTGAGCGGATAAAGCGCTCTG 2955  
Db 264352 ACCTGGTGGGCTCGGGGTCTGAAGCTGTGCAACCCCTTGAGCGGATAAAGCGCTCTG 264411  
QY 2956 TTTGCAACGGCGGCGTGAACGCGACCTGAACGGGACGCACTACACGTAACGGCGCGC 3015  
Db 264412 TTTGCAACGGCGGCGTGAACGCGACCTGAACGGGACGCACTACACGTAACGGCGCGC 264471

QY 3016 TTTTACCGCGCGACTGCAGCAACCGCGCAACGCGGGCAGCAATATGCCGACACCCGC 3075  
Db 264472 TTTTACCGCGCGACTGCAGCAACCGCGCAACGCGGGCAGCAATATGCCGACACCCGC 264531  
QY 3076 CTGCTTCCGCGGCTGCGCGCGGATGTCGAATTCGGAACGCGTGGACGCGTTGGCAGCT 3135  
Db 264532 CTGCTTCCGCGGCTGCGCGCGGATGTCGAATTCGGAACGCGTGGACGCGTTGGCAGCT 264591  
QY 3136 TACAGTACCGCGGTTCCAAACAGTACGGCAACCAACGCGGAGTTCGCGGTAGGCTAC 3195  
Db 264592 TACAGTACCGCGGTTCCAAACAGTACGGCAACCAACGCGGAGTTCGCGGTAGGCTAC 264651  
QY 3196 CGGTTCTGA 3204  
Db 264652 CGGTTCTGA 264660  
  
RESULT 4  
AAS43905  
ID AAS43905 standard; DNA; 4218 BP.  
XX AAS43905;  
XX  
DT 18-DEC-2001 (first entry)  
XX  
DE Neisseria meningitidis fusion protein 961cL-983 DNA.  
XX  
KW Neisseria gonorrhoeae; leader peptide; fusion protein; ORF46.1; ds;  
KW Neisserial protein.  
XX  
OS Neisseria meningitidis.  
OS Synthetic.  
PN WO200164922-A2.  
XX  
PD 07-SEP-2001.  
XX  
PF 28-FEB-2001; 2001WO-1B00452.  
XX  
PR 28-FEB-2000; 2000GB-0004695.  
PR 13-NOV-2000; 2000GB-0027675.  
XX  
PA (CHIR-) CHIRON SPA.  
PI Arico MB, Comanducci M, Galeotti C, Masignani V, Guilliani MM;  
PI Pizza M;  
XX  
XX WPI; 2001-582163/65.  
DR P-PSDB; AAU27608.  
XX  
PT Producing heterologous proteins from Neisseria meningitidis and N.  
PT gonorrhoeae -  
XX  
XX Example 23; Page 75-76; 119pp; English.  
PS  
XX The invention relates to methods for the heterologous expression of  
CC Neisserial proteins from Neisseria meningitidis and Neisseria  
CC gonorrhoeae. At least one domain in the protein is deleted, e.g. the  
CC leader peptide, and may be replaced by a domain from a different protein  
CC to make a fusion protein, in order to enhance heterologous expression of  
CC Neisserial proteins. Also, a region of a protein, such as a poly-glycine  
CC stretch, can be mutated to enhance expression. The proteins used in the  
CC processes include ORF46.1, 287, 741, 919, 953, 961 and 983. Sequences  
CC AAS43868-AAS43905 represent DNA molecules encoding Neisserial proteins  
CC and peptide regions of proteins of the invention.  
XX  
SQ Sequence 4218 BP; 1196 A; 1149 C; 1130 G; 743 T; 0 other;

Query Match 93.2%; Score 2986.2; DB 22; Length 4218;  
Best Local Similarity 97.1%; Pred. No. 0;  
Matches 3067; Conservative 0; Mismatches 48; Indels 42; Gaps 1;

QY 90 CGGCGGGCGGCACCTTCTGCGCCCGACCTTCAATGCAGGCGGCACCGGTATCGGCAGCAA 149  
|||||  
Db 1056 CGGCGAGGCGCAGCTTCTGCGCCGAGCTTCAATGCAGGCGGTACCGGTATCGGCAGCAA 1115  
QY 150 CAGCAGACACAACAGCGAATACGACGACGATCTTACGCGGTATCAAGAAAGAAAT 209  
|||||  
Db 1116 CAGCAGACACAACAGCGAATACGACGACGATCTTACGCGGTATCAAGAAAGAAAT 1175  
QY 210 GTGCAAGACAGAAGCATGCTCTGTGCGCGGTGGGATGACGTTGCGGTACAGACAGGA 269  
|||||  
Db 1176 GTGCAAGACAGAAGCATGCTCTGTGCGCGGTGGGATGACGTTGCGGTACAGACAGGA 1235  
QY 270 TGCCAAATCAATGCCCGCCCGCCGAAATCTGCATACCGGAGACTTTACAAACCCCAATGA 329  
|||||  
Db 1236 TGCCAAATCAATGCCCGCCCGCCGAAATCTGCATACCGGAGACTTTTCCAAACCCCAATGA 1295  
QY 330 CCGATACAGAATTTGATCAACCTCAACCTTCAACCTTGAAGCAGGCTATACAGGACGGG 389  
|||||  
Db 1296 CCGATACAGAATTTGATCAACCTCAACCTTCAACCTTGAAGCAGGCTATACAGGACGGG 1355  
QY 390 GGTAGAGGTAGGTATGCTGATACAGGCGAATCCGCTCGGCACGATATCCTTTCCCGAACT 449  
|||||  
Db 1356 GGTAGAGGTAGGTATGCTGATACAGGCGAATCCGCTCGGCACGATATCCTTTCCCGAACT 1415  
QY 450 GTATGGCAGAAAGAACACGCGCTATAACGAAATTTACAAAACTATACGGCGTATATGCG 509  
|||||  
Db 1416 GTATGGCAGAAAGAACACGCGCTATAACGAAATTTACAAAACTATACGGCGTATATGCG 1475  
QY 510 GAAGCAAGCGCTGAAGACGAGCGGTAAAGACATTTAAAGCTTCTTCGACGATGAGGC 569  
|||||  
Db 1476 GAAGCAAGCGCTGAAGACGAGCGGTAAAGACATTTAAAGCTTCTTCGACGATGAGGC 1535  
QY 570 CATTATAGAGACTGAAGCAAAAGCCGACGATATCCGCGCACGTAATGAAAGAAATCGGACACAT 629  
|||||  
Db 1536 CATTATAGAGACTGAAGCAAAAGCCGACGATATCCGCGCACGTAATGAAAGAAATCGGACACAT 1595  
QY 630 CGATGTGCTCCCATATTTATTTGGCGGGGTTCCGCTGACGCGCAGACCTGACGGCGGTAT 689  
|||||  
Db 1596 CGATTTGCTCTCCCATATTTATTTGGCGGGGTTCCGCTGACGCGCAGACCTGACGGCGGTAT 1655  
QY 690 TCGCGCGGATGGCGCTACACATAATGAATACGATGATGAACCAAGCAAGAAATAT 749  
|||||  
Db 1656 TCGCGCGGATGGCGCTACACATAATGAATACGATGATGAACCAAGCAAGAAATAT 1715  
QY 750 GTCTCAGCCATCCGCAATGATGGGTCAAGCTGGGCGAAACGCTGGCGCATCGTCAA 809  
|||||  
Db 1716 GGTTCAGCCATCCGCAATGATGGGTCAAGCTGGGCGAAACGCTGGCGCATCGTCAA 1775  
QY 810 TAAAGTTTTGAAACAACATCGAGGCGAGGCTCGCGACCAATTTCCAAATAGCCAAATTC 869  
|||||  
Db 1776 TAAAGTTTTGAAACAACATCGAGGCGAGGCACTGCGGACCTTTTCCAAATAGCCAAATTC 1835  
QY 870 GGAGGACAGTACCGCAGCGTTCTCGCTATTTCCGCGGTGATAAAGACAGAGGG 929  
|||||  
Db 1836 GGAGGACAGTACCGCAGCGTTCTCGCTATTTCCGCGGTGATAAAGACAGAGGG 1895  
QY 930 TATCCGCTGATGCAACAGAGGATTTAGGCAACTTGCTTACCACATCCGCTAATAAAAA 989  
|||||  
Db 1896 TATCCGCTGATGCAACAGAGGATTTAGGCAACCTGCTTACCACATCCGCTAATAAAAA 1955  
QY 990 CATGCTTTTCATTTTTCGGCAAGCAATGACGCAAGCTAGCCCAACACACTGACCT 1049  
|||||  
Db 1956 CATGCTTTTCATTTTTCGGCAAGCAATGACGCAAGCTAGCCCAACACACTGACCT 2015  
QY 1050 ATTGCAATTTTATGAAAGATGCTCAAAAAGGCAATTTACAGCTCGGCGGTAGACCG 1109  
|||||  
Db 2016 ATTGCAATTTTATGAAAGATGCTCAAAAAGGCAATTTACAGCTCGGCGGTAGACCG 2075  
QY 1110 CAGTGAGAAAAGTTCA -----A 1127  
|||||  
Db 2076 CAGTGAGAAAAGTTCAACAGGGAATGTATGGAAACCGGGTACAGAACCGCTTGAGTA 2135  
QY 1128 TGGCTCCAACCATTCGGGAATTTACTGCCATGTGTGCTATTCGGCACCCCTATGAAGCAAG 1187  
|||||

Db 2136 TGGCTCCAACCATTCGGGAATTTACTGCCATGTGTGCTGTGCGCACCTATGAAGCAAG 2195  
QY 1188 GGTCCGTTTCACCCGTACAAACCCGATTCAAATTCGCGGAACATCCTTTTCCGCAACCAT 1247  
Db 2196 GGTCCGTTTCACCCGTACAAACCCGATTCAAATTCGCGGAACATCCTTTTCCGCAACCAT 2255  
QY 1248 GGTAAACCGCAGCGGCGCTCTGCTCTGCAGAAATACCCGTGGATGAGCAACGACAACCT 1307  
|||||  
Db 2256 GGTAAACCGCAGCGGCGCTCTGCTCTGCAGAAATACCCGTGGATGAGCAACGACAACCT 2315  
QY 1308 GGTATACCAGCTGTGTGACAACAGGCTCAGACATCCGTTGCGAGTCGCGGTGAGACAGAAATT 1367  
|||||  
Db 2316 GGTATACCAGCTGTGTGAGACAGGCTCAGGACATCCGTTGCGAGTCGCGGTGAGACAGAAATT 2375  
QY 1368 GGTCTGGGAGCTGCTGGATGCGGGTAAAGCCATGAAGCCATGAAGCCCGCGCTCTTTCCGTTGG 1427  
Db 2376 GGTCTGGGAGCTGCTGGATGCGGGTAAAGCCATGAAGCCCGCGCTCTTTCCGTTGG 2435  
QY 1428 GGTCTTACCGCGGATACGAAAGGTACATCCGATTTGCTACTCTCTCCGTAAAGACAT 1487  
Db 2436 GGTCTTACCGCGGATACGAAAGGTACATCCGATTTGCTACTCTCTCCGTAAAGACAT 2495  
QY 1488 TTAGGCAACGCGCGCTGTATCAAAAAAGGCGCAGCAACTGCAACTGCACGGCAACAA 1547  
Db 2496 TTAGGCAACGCGCGCTGTATCAAAAAAGGCGCAGCAACTGCAACTGCACGGCAACAA 2555  
QY 1548 CACTTATACGGGCAAAACCATTTATCGAAGGCGTTTCGCTGGTGTAGCGCAACAA 1607  
Db 2556 CACTTATACGGGCAAAACCATTTATCGAAGGCGTTTCGCTGGTGTAGCGCAACAA 2615  
QY 1608 ATCGGATATGGCGCTCGAAACCAAGGTGCGCTGATTTATACGGGGGGGATCCTCGGGG 1667  
Db 2616 ATCGGATATGGCGCTCGAAACCAAGGTGCGCTGATTTATACGGGGGGGATCCTCGGGG 2675  
QY 1668 TAGCTGACAGCGACGCGATTTGCTATCTGCGACATACCGACCGATCGGCGCAACGA 1727  
Db 2676 CAGCTGACAGCGACGCGATTTGCTATCTGGCAGATACCGACCAATCCGCGCAACGA 2735  
QY 1728 AACCTGACATCAAAAGCGGATCTCAGTGGCGGGAAGGTACGCTGTACACAGCTTT 1787  
Db 2736 AACCTGACATCAAAAGCGGATCTCAGTGGCGGGAAGGTACGCTGTACACAGCTTT 2795  
QY 1788 GGGCAAACTGCTGAAAGTGGACGCTACGCGGATGACCGCGCAAGCTGTACATGTCGC 1847  
Db 2796 GGGCAAACTGCTGAAAGTGGACGCTACGCGGATTTATCGCGCAAGCTGTACATGTCGC 2855  
QY 1848 AGCGGCAAAAGGCGGCGATCTCAACGCTACCGCAACAGTGTTCCTTCCCTGAGTGC 1907  
Db 2856 AGCGGCAAAAGGCGGCGATCTCAACAGTACCGGACGCTGTTCCTTCCCTGAGTGC 2915  
QY 1908 GGGCAAAATCGGCGGATTTCTTCTTCAAAACATCGAAACCGACGGTGTCTGCT 1967  
Db 2916 GGGCAAAATCGGCGGATTTCTTCTTCAAAACATCGAAACCGACGGCGGCTGCT 2975  
QY 1968 GGTTCCTTCGACAGCGCTCGAAAAAACAGCGGCGAGTGAAGGCGACAGCTGTCTATT 2027  
Db 2976 GGTTCCTTCGACAGCGCTCGAAAAAACAGCGGCGAGTGAAGGCGACAGCTGTCTATT 3035  
QY 2028 TGTCCGTCCGCGCAATCGGCGACGAGCTGCTTCGCGACGGGCACATTCGCGCGCGCGG 2087  
Db 3036 TGTCCGTCCGCGCAATCGGCGACGAGCTGCTTCGCGACGGGCACATTCGCGCGCGCGG 3095  
QY 2088 TGTGAAACACGCGGTAGAACAGGCGGCGAGCAATCTGAAACACCTGATGGTGAACCTGGA 2147  
Db 3096 TGTGAAACACGCGGTAGAACAGGCGGCGAGCAATCTGAAACACCTGATGGTGAACCTGGA 3155  
QY 2148 TGTCTCCGAATCATCCGCAACACCGGAGACGCTTGAACCTGCGCGCGCGGACGACAGA 2207  
Db 3156 TGTCTCCGAATCATCCGCAACACCGGAGACGCTTGAACCTGCGCGCGCGGACGACAGA 3215  
QY 2208 TATGCGCGGCGATCCGCCCTACGGCGCAACTTTCGCGCGAGCGGCGGCTACAGCATGC 2267  
|||||

Df 3216 TATGCGGGCATCCGCCCTACGGGCAACTTCCGCGCAGCGGCGATACAGCATGC 3275  
QY 2268 GAATGCCGCGAGCGGTGTACGATCTTCAACAGTCTTCGCGCTACCGTCTATGCCGACAG 2327  
Df 3276 GAATGCCGCGAGCGGTGTACGATCTTCAACAGTCTTCGCGCTACCGTCTATGCCGACAG 3335  
QY 2328 TACCGCGGCCCATGCGGATATGACAGGACGCGCGGTGAAGACCGTATCGGACGGGTGGA 2387  
Df 3336 TACCGCGGCCCATGCGGATATGACAGGACGCGCGGTGAAGACCGTATCGGACGGGTGGA 3395  
QY 2388 CCACACGCTAGCGGTCTCGCGGTATCGCGCAACCCCAACAGGACGGTGAACGTGGA 2447  
Df 3396 CCACACGCGCGGTCTCGCGGTATCGCGCAACCCCAACAGGACGGTGAACGTGGA 3455  
QY 2448 ACAGGCGGTGTGAAGGCAAAATGCGGCGAGTACCCAAACCGTCGGCATTCGCCGGA 2507  
Df 3456 ACAGGCGGTGTGAAGGCAAAATGCGGCGAGTACCCAAACCGTCGGCATTCGCCGGA 3515  
QY 2508 AACCGCGAAATAGCACAGCGCGCCACACTGGGCATGGGACACAGCACATGGAGCGA 2567  
Df 3516 AACCGCGAAATAGCACAGCGCGCCACACTGGGCATGGGACACAGCACATGGAGCGA 3575  
QY 2568 AAACAGTGAATGCAAAACGACAGCATTAAGTCTGTTTTCAGGACATACGGACGATGC 2627  
Df 3576 AAACAGTGAATGCAAAACGACAGCATTAAGTCTGTTTTCAGGACATACGGACGATGC 3635  
QY 2628 GGGCATATCGGCTATCTCAAGGCGTGTCTCTACGAGCGCTACAAAACAGCATCAG 2687  
Df 3636 GGGCATATCGGCTATCTCAAGGCGTGTCTCTACGAGCGCTACAAAACAGCATCAG 3695  
QY 2688 CGCGACACCGGTGGCGGAGCAATGCGGAAGCGAGCGTCAACGGGACGCTGATGAGCT 2747  
Df 3696 CGCGACACCGGTGGCGGAGCAATGCGGAAGCGAGCGTCAACGGGACGCTGATGAGCT 3755  
QY 2748 GGGCGACTGGCGGTGTCAAGTTCGGTTCGCGCAACGGGAGATTTGACGGTCAAGG 2807  
Df 3756 GGGCGACTGGCGGTGTCAAGTTCGGTTCGCGCAACGGGAGATTTGACGGTCAAGG 3815  
QY 2808 CGGTCTCGCTACGACTCTCAACAGGATGATTCGCGGCAAAAGGAGTCTTGGG 2867  
Df 3816 CGGTCTCGCTACGACTCTCAACAGGATGATTCGCGGCAAAAGGAGTCTTGGG 3875  
QY 2868 CTGAGCGGCAACAGCCTCACTGAAGGACACTGCTGCGACTCGCGGTCTGAAGCTGC 2927  
Df 3876 CTGAGCGGCAACAGCCTCACTGAAGGACACTGCTGCGACTCGCGGTCTGAAGCTGC 3935  
QY 2928 GCAACCTTTGAGCGATAAAGCGTCTCTGTTTGAACGGCGGGGTGGAACGGACCTGAA 2987  
Df 3936 GCAACCTTTGAGCGATAAAGCGTCTCTGTTTGAACGGCGGGGTGGAACGGACCTGAA 3995  
QY 2988 CGGACGCGACTACCGGTAACGGCGGCTTTACCGGCGGACTGCAAGCAACGGCAAGAC 3047  
Df 3996 CGGACGCGACTACCGGTAACGGCGGCTTTACCGGCGGACTGCAAGCAACGGCAAGAC 4055  
QY 3048 GGGGCGACCAATATGCCGCACACCGCGCTGTTGCGGCTCTGGGCGGATGTCGAATT 3107  
Df 4056 GGGGCGACCAATATGCCGCACACCGCGCTGTTGCGGCTCTGGGCGGATGTCGAATT 4115  
QY 3108 CGGCAACGCTGGAACGGCTTGGCAGTTACAGCTACGCGGTTTCCAAACAGTACGGCAA 3167  
Df 4116 CGGCAACGCTGGAACGGCTTGGCAGTTACAGCTACGCGGTTTCCAAACAGTACGGCAA 4175  
QY 3168 CCACAGCGGACGAGTGGCGGTAGGCTACCGGTTCTGA 3204  
Df 4176 CCACAGCGGACGAGTGGCGGTAGGCTACCGGTTCTGA 4212

RESULT 5

AD17058  
ID AD17058 standard; DNA; 4218 BP.

XX

AC AD17058;

XX

DT 29-NOV-2001 (first entry)  
XX N. meningitidis strain 2996 961cL-983 fusion DNA.  
DE Heterologous expression; Neisserial protein;  
KW 961cL-983 fusion protein; ds.  
XX Neisseria meningitidis 2996.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..4212  
FT /\*tag= a  
FT /product= "N. meningitidis strain 2996 961cL-983  
FT fusion protein"  
XX  
PN WO200164920-A2.  
XX  
PD 07-SEP-2001.  
XX  
PF 28-FEB-2001; 2001WO-IB00420.  
XX  
PR 28-FEB-2000; 2000GB-0004695.  
PR 13-NOV-2000; 2000GB-0027675.  
XX  
PA (CHIR-) CHIRON SPA.  
XX  
PI Arico MB, Comanducci M, Galeotti C, Masignani V, Giuliani MM;  
PI Pizza M;  
XX  
DR WPI; 2001-557776/62.  
DR P-PSDB; AAE10043.  
XX  
PT Heterologous expression for the expression of two or more Neisserial  
PT proteins in fused state  
XX  
PS Example 23; Page 33-34; 52pp; English.  
XX  
CC The present invention relates to a method for simultaneous heterologous  
CC expression of two or more Neisserial proteins which are in a fused  
CC state. The method is useful for simultaneous heterologous expression of  
CC two or more Neisserial proteins. A protein that may be unstable or  
CC poorly expressed on its own is assisted by adding a suitable hybrid  
CC partner and commercial manufacture is simplified-only one expression and  
CC purification need to be employed in order to produce two separately-  
CC useful proteins. The present sequence is a DNA encoding  
CC Neisseria meningitidis (serogroup B, strain 2996) 961cL-983  
CC fusion protein.  
XX  
SQ Sequence 4218 BP; 1196 A; 1149 C; 1130 G; 743 T; 0 other;  
QY 90 CGCGCGCGCGGCACATTCGCGCCGACTTCAATGAGGCGGCGGATCGGAGCAA 149  
Df 1056 CGCGCGCGCGGCACATTCGCGCCGACTTCAATGAGGCGGATCGGAGCAA 1115  
QY 150 CAGCAGAGCAACAGCGAAATCAGCAGAGTATCTTACGCGGATCAAGAACGAAT 209  
Df 1116 CAGCAGAGCAACAGCGAAATCAGCAGAGTATCTTACGCGGATCAAGAACGAAT 1175  
QY 210 GTGCAAGACAGAGCATGCTCTGTCGCGGTGGGATGAGTTTACACAGGGA 269  
Df 1176 GTGCAAGACAGAGCATGCTCTGTCGCGGTGGGATGAGTTTACACAGGGA 1235  
QY 270 TGCCAAATCAATGCCCGCCCGCCCGAATCTGCATACCGGAGACTTTACAAACCAATGA 329  
Df 1236 TGCCAAATCAATGCCCGCCCGCCCGAATCTGCATACCGGAGACTTTCCAAACCAATGA 1295  
QY 330 CGCATACAGAAATTTGATCAACCTCAACCTGCAATTTGAAGCGGTATACAGGACGG 389  
Df 1296 CGCATACAGAAATTTGATCAACCTCAACCTGCAATTTGAAGCGGTATACAGGACGG 1355



Db 3516 AACCGCGAAAAATACACAGCAGCCGCCACACTGGCGATGGAGCAGCACATGGAGCGA 3575  
QY 2568 AAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGGAGGCATACGGCACCAGTC 2627  
Db 3576 AAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGGAGGCATACGGCACCAGTC 3635  
QY 2628 GGGCGATATCGGCTATCTCAAAAGCGCTGTCTCTACGGACGCTACAAAAACAGCATCAG 2687  
Db 3636 GGGCGATATCGGCTATCTCAAAAGCGCTGTCTCTACGGACGCTACAAAAACAGCATCAG 3695  
QY 2688 CCGCAGCACCCTGGCGGAGCAGACATCGGAAGCAGCGTCAACGGCAGCGCTGATGACGCT 2747  
Db 3696 CCGCAGCACCCTGGCGGAGCAGACATCGGAAGCAGCGTCAACGGCAGCGCTGATGACGCT 3755  
QY 2748 GGGCGCAGTGGCGGCTGTCAGCTTCCGTTTCCGCAACGGGAGATTTACGGCTGCAAGG 2807  
Db 3756 GGGCGCAGTGGCGGCTGTCAGCTTCCGTTTCCGCAACGGGAGATTTACGGCTGCAAGG 3815  
QY 2808 CGGTCTGCGCTACGAGCTGCTCAACAGGATGATTCGCGCAAAAGGAGCGTCTTTGGG 2867  
Db 3816 CGGTCTGCGCTACGAGCTGCTCAACAGGATGATTCGCGCAAAAGGAGCGTCTTTGGG 3875  
QY 2868 CTGGAGCGGCAACAGCCTCACTGAAGGCACATGTGTGCGACTGCGGGTCTGAAGCTGTC 2927  
Db 3876 CTGGAGCGGCAACAGCCTCACTGAAGGCACATGTGTGCGACTGCGGGTCTGAAGCTGTC 3935  
QY 2928 GCAACCCCTTGAGCGATAAAGCGTCTCTGTTTCAACGGCGGGCTGGAACGCGACCTGAA 2987  
Db 3936 GCAACCCCTTGAGCGATAAAGCGTCTCTGTTTCAACGGCGGGCTGGAACGCGACCTGAA 3995  
QY 2988 CGGACGCGACTACAGCTTAACGGCGGCTTTACCGCGCGACTGCAAGCAACCGGCAAGAC 3047  
Db 3996 CGGACGCGACTACAGCTTAACGGCGGCTTTACCGCGCGACTGCAAGCAACCGGCAAGAC 4055  
QY 3048 GGGGGCACCATAATATCCCGCACACCCGCTGTTGGTGGCGGTCTGGGCGCGGATGTCGAATT 3107  
Db 4056 GGGGGCACCATAATATCCCGCACACCCGCTGTTGGTGGCGGTCTGGGCGCGGATGTCGAATT 4115  
QY 3108 CGGCAACGGCTGGAACGGCTTGGCAGCTTACAGCTACGCGCGGTTCCAAACAGTACGCGCAA 3167  
Db 4116 CGGCAACGGCTGGAACGGCTTGGCAGCTTACAGCTACGCGCGGTTCCAAACAGTACGCGCAA 4175  
QY 3168 CCACAGCGAGCTGCGGCTAGGCTACCGGTTCTGA 3204  
Db 4176 CCACAGCGAGCTGCGGCTAGGCTACCGGTTCTGA 4212

RESULT 6  
AAS43899  
ID AAS43899 standard; DNA; 4335 BP.  
AC AAS43899;  
XX AAS43899;  
AC AAS43899;  
DT 18-DEC-2001 (first entry)  
XX Neisseria meningitidis fusion protein 961-983 DNA.  
DE Neisseria meningitidis fusion protein; ORF46.1; ds;  
KW Neisseria gonorrhoeae; leader peptide; fusion protein; ORF46.1; ds;  
KW Neisserial protein.  
XX Neisseria meningitidis.  
OS Neisseria meningitidis.  
XX WO200164922-A2.  
PN 07-SEP-2001.  
XX 28-FEB-2001; 2001WO-IB00452.  
PF 28-FEB-2000; 2000GB-0004695.  
PR 13-NOV-2000; 2000GB-002765.  
XX

(CHIR-) CHIRON SPA.  
XX Arico MB, Comanducci M, Galeotti C, Masignani V, Guilliani MM;  
PI Pizza M;  
XX WPI; 2001-582163/65.  
DR P-PSDB; AAU27602.  
XX Producing heterologous proteins from Neisseria meningitidis and N.  
gonorrhoeae -  
XX Example 23; Page 68-69; 119pp; English.  
XX The invention relates to methods for the heterologous expression of  
CC Neisserial proteins from Neisseria meningitidis and Neisseria  
CC gonorrhoeae. At least one domain in the protein is deleted, e.g. the  
CC leader peptide, and may be replaced by a domain from a different protein  
CC to make a fusion protein, in order to enhance heterologous expression of  
CC Neisserial proteins. Also, a region of a protein, such as a poly-glycine  
CC stretch, can be mutated to enhance expression. The proteins used in the  
CC processes include ORF46.1, 287, 741, 919, 953, 961 and 983. Sequences  
CC AAS43889-AAS43905 represent DNA molecules encoding Neisserial proteins  
CC and peptide regions of proteins of the invention.  
XX Sequence 4335 BP; 1217 A; 1189 C; 1165 G; 764 T; 0 other;  
SQ  
Query Match 93.1%; Score 2983.8; DB 22; Length 4335;  
Best Local Similarity 96.9%; Pred. No. 0;  
Matches 3070; Conservative 0; Mismatches 57; Indels 42; Gaps 1;  
QY 75 TTCTGCTGCTTAGCGGGCGGCGCACTTCTGCGCCGCACTTCAATCGAGCGGCAC 134  
Db 1140 TTACGAGTGGGATCGGGAGGCGGCATCTGCGCCGACTTCAATCGAGCGGTAC 1199  
QY 135 CGGTATCGGAGCAACAGCAGAGCAACAACAGCGAAATCAGCAGTAGTATCTTACGCCGG 194  
Db 1200 CGGTATCGGAGCAACAGCAGAGCAACAACAGCGAAATCAGCAGTAGTATCTTACGCCGG 1259  
QY 195 TATCAAGACGAATCTGCAAAAGACAGAGCATGCTCTGCGCGGTCGGATCACGCTTC 254  
Db 1260 TATCAAGACGAATCTGCAAAAGACAGAGCATGCTCTGCGCGGTCGGATCACGCTTC 1319  
QY 255 GGTACAGACAGGATGCCAAATCAATGCCCGCCCGCAATCTGCATACCGGAGACTT 314  
Db 1320 GGTACAGACAGGATGCCAAATCAATGCCCGCCCGCAATCTGCATACCGGAGACTT 1379  
QY 315 TACAAACCCAAATGACGCGATACAAGAAATTTGATCAACCTCAAACCTGCAATTTGAACGAGG 374  
Db 1380 TCCAAACCCAAATGACGCGATACAAGAAATTTGATCAACCTCAAACCTGCAATTTGAACGAGG 1439  
QY 375 CTATACAGGACGGGTAGAGTAGGTATCGTCGATACAGGCGAATCGCTCGGCGACAT 434  
Db 1440 CTATACAGGACGGGTAGAGTAGGTATCGTCGACACAGGCGAATCGCTCGGCGACAT 1499  
QY 435 ATCCTTTCCGCACTGTATGGCAGAAAAAAGAACACGCGTATACGAAAAATACAAAACTA 494  
Db 1500 ATCCTTTCCGCACTGTATGGCAGAAAAAAGAACACGCGTATACGAAAAATACAAAACTA 1559  
QY 495 TAGCGGTATATGCGGAAGAAAGCGCTGAAGACGAGGCGGTAAGACATTTAAAGCTTC 554  
Db 1560 TAGCGGTATATGCGGAAGAAAGCGCTGAAGACGAGGCGGTAAGACATTTAAAGCTTC 1619  
QY 555 TTTCGACGATGAGCGGTTATAGAGACTGAGCAAAAGCGACGGATATCGGCCACGTAA 614  
Db 1620 TTTCGACGATGAGCGGTTATAGAGACTGAGCAAAAGCGACGGATATCGGCCACGTAA 1679  
QY 615 AGAAATCGGACACATCGATGTGCTCCCATATTTATGGCGGCGTTCGCTGACCGGCGAG 674  
Db 1680 AGAAATCGGACACATCGATTTGCTCCCATATTTATGGCGGCGTTCGCTGACCGGCGAG 1739  
QY 675 ACCTGCAGCGGTATTTGCGCCCGATCGAGCTACACATAATGAATACGATGATGAAC 734  
Db 1740 ACCTGCAGCGGTATTTGCGCCCGATCGAGCTACACATAATGAATACGATGATGAAC 1799

QY 735 CAAGAACGAAATAAATGTTCTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAAGCTGG 794  
DB 1800 CAAGAAGCAATGATGGTTGCGACCATCCGCAATGCATGGGTCAAGCTGGGCGAAGCTGG 1859  
QY 795 CGTGGCATCGTCAATACAGTTTGGAAACACATCGAGGCGAGGACACTGCGGACCATTT 854  
DB 1860 CGTGGCATCGTCAATACAGTTTGGAAACACATCGAGGCGAGGACACTGCGGACCATTT 1919  
QY 855 CCAATAGCAATTCGGAGGACGATACCGCAAGCTTGTCTGCCTATTTCGGCGGTGA 914  
DB 1920 CCAATAGCAATTCGGAGGACGATACCGCAAGCTTGTCTGCCTATTTCGGCGGTGA 1979  
QY 915 TAAACAGACAGGATTCGGCTGATCAACAGAGCGATTACGGCAACTTGTCTCAACCA 974  
DB 1980 TAAACAGACAGGATTCGGCTGATCAACAGAGCGATTACGGCAACTTGTCTCAACCA 2039  
QY 975 CATCCGTAAATAACATGCTTTTCATTTTTCGGCAAGCAATGACGCAACAGCTCAGCC 1034  
DB 2040 CATCCGTAAATAACATGCTTTTCATTTTTCGGCAAGCAATGACGCAACAGCTCAGCC 2099  
QY 1035 CAACACACTGACCCATTATGCCATTTTATGAAAAGATGCTCAAAAAGCATTTATCACAGT 1094  
DB 2100 CAACACACTGACCCATTATGCCATTTTATGAAAAGATGCTCAAAAAGCATTTATCACAGT 2159  
QY 1095 CGCAGCGGTAGACCGAGTGGAGAAAAGTTCA - - - - - 1126  
DB 2160 CGCAGCGGTAGACCGAGTGGAGAAAAGTTCAACGGGAAATGTATGGAGAACCGGTATC 2219  
QY 1127 - - - - - NYGGCTCAACCATTTGCGGAATTAATGTCCTATGTTGGCTATFCGCG 1172  
DB 2220 AGAAGCGGTTGAGTATGGCTTCAACCATTTGCGGAATTAATGTCCTATGTTGGCTATFCGCG 2279  
QY 1173 ACCCTATGAAGCAAGCTGCGTTTACCCGTCACACCGATTCAAAATGCGCGAACATC 1232  
DB 2280 ACCCTATGAAGCAAGCTGCGTTTACCCGTCACACCGATTCAAAATGCGCGAACATC 2339  
QY 1233 CTTTTCGCGACCCATCGTAACCGGACGCGGCTCTGTGTCGAGAAATACCCCTGGAT 1292  
DB 2340 CTTTTCGCGACCCATCGTAACCGGACGCGGCTCTGTGTCGAGAAATACCCCTGGAT 2399  
QY 1293 GAGCAAGCAACCTGCGTACCAACCGCTGCTGACACGGCTCAGGACATCGGTGACGTGG 1352  
DB 2400 GAGCAAGCAACCTGCGTACCAACCGCTGCTGACACGGCTCAGGACATCGGTGACGTGG 2459  
QY 1353 CGTGACAGCAAGTTGCGCTGGGAGCTGCTGATCGGGTAAGGCCATGACCGGACCCGC 1412  
DB 2460 CGTGACAGCAAGTTGCGCTGGGAGCTGCTGATCGGGTAAGGCCATGACCGGACCCGC 2519  
QY 1413 GTCCCTTCGCTTCGGGACTTTTACCGCGGATACGAAAGGTACATCCGATATTGCGCTACTC 1472  
DB 2520 GTCCCTTCGCTTCGGGACTTTTACCGCGGATACGAAAGGTACATCCGATATTGCGCTACTC 2579  
QY 1473 CTTTCGTAACGACATTTACGGCACGGGGGCTGATCAAAAAGGCGGAGCCAACTGCA 1532  
DB 2580 CTTTCGTAACGACATTTACGGCACGGGGGCTGATCAAAAAGGCGGAGCCAACTGCA 2639  
QY 1533 ACTGCACGCAACAACTATACGGGCAAAACCATTTATCGAAGCGGTTGCTGCTGTT 1592  
DB 2640 ACTGCACGCAACAACTATACGGGCAAAACCATTTATCGAAGCGGTTGCTGCTGTT 2699  
QY 1593 GTACGGCAACAACTATCGGCTCGAAACCAAGGTGCGCTGATTATTAACGG 1652  
DB 2700 GTACGGCAACAACTATCGGCTCGAAACCAAGGTGCGCTGATTATTAACGG 2759  
QY 1653 GCGCGCATCGCGGTAGCTGAACAGGACGGCATTTGTCTATCTGCGAGATACCGACCG 1712  
DB 2760 GCGCGCATCGCGGTAGCTGAACAGGACGGCATTTGTCTATCTGCGAGATACCGACCA 2819  
QY 1713 ATCCGCGCAACAGCAACCGTGCATCAAGGCGCATCTGCAAGCTGGGCGGAGAGGTAC 1772  
DB 2820 ATCCGCGCAACAGCAACCGTGCATCAAGGCGCATCTGCAAGCTGGGCGGAGAGGTAC 2879

QY 1773 GCTGTACACACGTTTGGCGAAACTGCTGAAAGTGGACGCTAGCGGATACACCGCGGCAA 1832  
DB 2880 GCTGTACACACGTTTGGCGAAACTGCTGAAAGTGGACGCTAGCGGATATACCGCGGCAA 2939  
QY 1833 GCTGTACATGTCGGCACGCGGCAAAAGGCGAGGCTATCTCAACCGTACCGGACACGTTGT 1892  
DB 2940 GCTGTACATGTCGGCACGCGGCAAAAGGCGAGGCTATCTCAACAGTACCGGACGAGTTGT 2999  
QY 1893 TCCCTTCTGAGTGGCGCAAAATCGGCGGATTTCTTTTTCACAAAACATCGAAAC 1952  
DB 3000 TCCCTTCTGAGTGGCGCAAAATCGGCGAGGATTTCTTTTTCACAAAACATCGAAAC 3059  
QY 1953 CGACGGTGTCTGTCTGCTTCCCTCGACAGCGTCAAAAACAGCGGCGAGTGAAGGCGA 2012  
DB 3060 CGACGGCGCTGCTGTCTTCCCTCGACAGCGTCAAAAACAGCGGCGAGTGAAGGCGA 3119  
QY 2013 CACGCTGTCTTATTTATGTCCTCGCGGCAATCGCGCAGGACTGCTTCGGCAGCGGCGA 2072  
DB 3120 CACGCTGTCTTATTTATGTCCTCGCGGCAATCGCGCAGGACTGCTTCGGCAGCGGCGA 3179  
QY 2073 TTCGCGCGCGCGCTGTGAACACGCGCTAGAACAGGCGGCGACCAATCTGGAACACT 2132  
DB 3180 TTCGCGCGCGCGCTGTGAACACGCGCTAGAACAGGCGGCGACCAATCTGGAACACT 3239  
QY 2133 GATGTCGAACCTGGATGCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGCG 2192  
DB 3240 GATGTCGAACCTGGATGCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGCG 3299  
QY 2193 CGCGACCGCAGATATGCGGGCATCCGCCCTACGGCGCAACTTTCGCGCGCAGCGCG 2252  
DB 3300 AGCGACCGCAGATATGCGGGCATCCGCCCTACGGCGCAACTTTCGCGCGCAGCGCG 3359  
QY 2253 AGCGTACAGCATGCGAATGCGCGCAGCGTGTACGCAATCTTCAACAGTCTCGCGCTAC 2312  
DB 3360 AGCGTACAGCATGCGAATGCGCGCAGCGTGTACGCAATCTTCAACAGTCTCGCGCTAC 3419  
QY 2313 CGTCTATCCGACAGTACCGCGCCCATCGCGATATGAGGGACGCCGCTGAAAGCGGT 2372  
DB 3420 CGTCTATCCGACAGTACCGCGCCCATCGCGATATGAGGGACGCCGCTGAAAGCGGT 3479  
QY 2373 ATCGACCGGTTGGACACACAGCTACGGGTCTCGCGCTCATCGCGCAACCCACACAGA 2432  
DB 3480 ATCGACCGGTTGGACACACAGCTACGGGTCTCGCGCTCATCGCGCAACCCACACAGA 3539  
QY 2433 CGGTGGAACGTGGGAACAGGCGGTGTTGAGGCAAAATGCGCGCAGTACCCCAACCGT 2492  
DB 3540 CGGTGGAACGTGGGAACAGGCGGTGTTGAGGCAAAATGCGCGCAGTACCCCAACCGT 3599  
QY 2493 CGGCATTCCCGGAAACCGCGGAAATACGACAGCAGCGCCACACTGGGCGATGGGACA 2552  
DB 3600 CGGCATTCCCGGAAACCGCGGAAATACGACAGCAGCGCCACACTGGGCGATGGGAGC 3659  
QY 2553 CAGCACATGGAGCGGAAACAGTGCAGAAATGCAAAAACCCACAGCATTTAGTCTGTTGCGAG 2612  
DB 3660 CAGCACATGGAGCGGAAACAGTGCAGAAATGCAAAAACCCACAGCATTTAGTCTGTTGCGAG 3719  
QY 2613 CATAGGCGACGATCGGGCGATATCGGCTATCTCAAGGCGCTGTTCTCTACGAGCGCTA 2672  
DB 3720 CATAGGCGACGATCGGGCGATATCGGCTATCTCAAGGCGCTGTTCTCTACGAGCGCTA 3779  
QY 2673 CAAAACAGCATCAGCCGACGACCGCGGAGCAACATGCGGAAGGACGCGTCAACGG 2732  
DB 3780 CAAAACAGCATCAGCCGACGACCGCGGAGCAACATGCGGAAGGACGCGTCAACGG 3839  
QY 2733 CACGCTGATGACGTGGCGGACACTGGGCGGTGTAACGTTTCGTTTTCGCGCAACGGGAGA 2792  
DB 3840 CACGCTGATGACGTGGCGGACACTGGGCGGTGTAACGTTTCGTTTTCGCGCAACGGGAGA 3899  
QY 2793 TTTGACGCTGCAAGCGGTCTGCGCTACGACTGCTCAAAACAGGATGATTCGCGCGAAAA 2852  
DB 3900 TTTGACGCTGCAAGCGGTCTGCGCTACGACTGCTCAAAACAGGATGATTCGCGCGAAAA 3959  
QY 2853 AGGAGTGTCTTTGGGCTGGAGCGGCAACAGGCTCACTGAAAGGCACACTGCTGCGGACTCGC 2912

|||||  
Db 3960 AGCAGTGTCTTGGGTGGAGCGGCACACGCCCTCACTGAAGGCACGCTGGTGGACTCGC 4019  
QY 2913 GGGTCTGAAGCTCTGCAACCCCTTGAGCGGATAAAGCCGCTCTCTTTGCAACGCGGCGT 2972  
Db 4020 GGGTCTGAAGCTCTGCAACCCCTTGAGCGGATAAAGCCGCTCTCTTTGCAACGCGGCGT 4079  
QY 2973 GGAACCGACCTGAACGGACGCGACTACACGCTAAACGGGCGGCTTTACCGGCGCGACTGC 3032  
Db 4080 GGAACCGACCTGAACGGACGCGACTACACGCTAAACGGGCGGCTTTACCGGCGCGACTGC 4139  
QY 3033 AGCAACCGCAACAGCGGGGCGAGCAATATGCCGCGACACCCCGCTGGTTGCCGCTGGG 3092  
Db 4140 AGCAACCGCAACAGCGGGGCGAGCAATATGCCGCGACACCCCGCTGGTTGCCGCTGGG 4199  
QY 3093 CGCGGATGTCGAATTCGGCAACGGCTGGAACGCTTGGACAGCTTACAGCTACCGCGGTTTC 3152  
Db 4200 CGCGGATGTCGAATTCGGCAACGGCTGGAACGCTTGGACAGCTTACAGCTACCGCGGTTTC 4259  
QY 3153 CAAACAGTAGCGGCAACACAGCGGAGGTCGGCGTAGGCTACCGGTTTC 3201  
Db 4260 CAAACAGTAGCGGCAACACAGCGGAGGTCGGCGTAGGCTACCGGTTTC 4308  
  
RESULT 7  
AAD17052  
ID AAD17052 standard; DNA; 4335 BP.  
AC AAD17052;  
XX  
DT 29-NOV-2001 (first entry)  
XX  
DE N. meningitidis strain 2996 961-983 fusion DNA.  
XX  
KW Heterologous expression; Neisserial protein; 961-983 fusion protein; ds.  
XX  
OS Neisseria meningitidis 2996.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..4335  
FT /tag= a  
FT /product= "N. meningitidis strain 2996 961-983  
fusion protein"  
XX  
PN WO200164920-A2.  
XX  
PD 07-SEP-2001.  
XX  
PF 28-FEB-2001; 2001WO-IB00420.  
XX  
PR 28-FEB-2000; 2000GB-0004695.  
PR 13-NOV-2000; 2000GB-0027575.  
XX  
PA (CHIR-) CHIRON SPA.  
XX  
PI Arico MB, Comanducci M, Galeotti C, Masignani V, Giuliani MM;  
PI Pizza M;  
XX  
DR WPI: 2001-557776/62.  
DR P-PSDB; AAEL0037.  
XX  
PT Heterologous expression for the expression of two or more Neisserial  
PT proteins in fused state  
XX  
PS Example 23; Page 28-29; 52pp; English.  
XX  
CC The present invention relates to a method for simultaneous heterologous  
CC expression of two or more Neisserial proteins which are in a fused  
CC state. The method is useful for simultaneous heterologous expression of  
CC two or more Neisserial proteins. A protein that may be unstable or  
CC poorly expressed on its own is assisted by adding a suitable hybrid  
CC partner and commercial manufacture is simplified-only one expression and  
CC purification need to be employed in order to produce two separately-

useful proteins. The present sequence is a DNA encoding  
CC Neisseria meningitidis (serogroup B, strain 2996) 961-983 fusion  
CC protein.  
XX  
SQ Sequence 4335 BP; 1217 A; 1189 C; 1165 G; 764 T; 0 other;  
  
Query Match 93.1%; Score 2983.8; DB 22; Length 4335;  
Best Local Similarity 96.9%; Pred. No. 0;  
Matches 3070; Conservative 0; Mismatches 57; Indels 42; Gaps 1;  
  
QY 75 TTCTGCTGCTTAGGCGGCGGCGGCACTTCTGCGCGCGACTTCAATCGACGGCGCAC 134  
Db 1140 TTACGATGGGGATCGCGGAGGCGGCACTTCTGCGCGCGACTTCAATCGACGGCGTAC 1199  
QY 135 CGTATTCGGCAGCAACAGCAGACAGCAACAACAGCGAAATCAGCAGCATGATCTTTACGCGG 194  
Db 1200 CGTATTCGGCAGCAACAGCAGACAGCAACAACAGCGAAATCAGCAGCATGATCTTTACGCGG 1259  
QY 195 TATCAAGACGAATGTGCAAGACAGACATGCTCTGTGCGCGGTGGGATGACGTTGC 254  
Db 1260 TATCAAGACGAATGTGCAAGACAGACATGCTCTGTGCGCGGTGGGATGACGTTGC 1319  
QY 255 GGTTCAGACAGGATGCCAAATCAATGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 314  
Db 1320 GGTTCAGACAGGATGCCAAATCAATGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1379  
QY 315 TACAAACCCAAATGACGCATACAGAAATTTGATCAACCTCAAACTTCAATTTGAAGCAGG 374  
Db 1380 TCCAAACCCAAATGACGCATACAGAAATTTGATCAACCTCAAACTTCAATTTGAAGCAGG 1439  
QY 375 CTATACAGGACGCGGGTAGAGTAGGTATCGTCGATACAGGCGAAATCCGTCGGCAGCAT 434  
Db 1440 CTATACAGGACGCGGGTAGAGTAGGTATCGTCGATACAGGCGAAATCCGTCGGCAGCAT 1499  
QY 435 ATCCTTTCCCGAATGTATGGCAGAAAGAACAGCGGTATTAACGAAATTTACAAAACATA 494  
Db 1500 ATCCTTTCCCGAATGTATGGCAGAAAGAACAGCGGTATTAACGAAATTTACAAAACATA 1559  
QY 495 TACGGCGTATATCGGAAGGAGCGCTGAAGACGAGGCGGTAAAGACATTTAAAGGTTTC 554  
Db 1560 TACGGCGTATATCGGAAGGAGCGCTGAAGACGAGGCGGTAAAGACATTTAAAGGTTTC 1619  
QY 555 TTTCGACGATGAGCGCGTTATAGAGACTGAAGCAAAAGCCGACGATATCCGCCACGTA 614  
Db 1620 TTTCGACGATGAGCGCGTTATAGAGACTGAAGCAAAAGCCGACGATATCCGCCACGTA 1679  
QY 615 AGAAATCGGACACATCGATGTGCTCCCATATTTATTTGGCGGGCGGTTCCGTGACGGCAG 674  
Db 1680 AGAAATCGGACACATCGATTTGGTCTCCCATATTTATTTGGCGGGCGGTTCCGTGACGGCAG 1739  
QY 675 ACCTGCAGCGGTATTGCGCGCGATCGACGCTACACATATGAATACGATGATGAAC 734  
Db 1740 ACCTGCAGCGGTATTGCGCGCGATCGACGCTACACATATGAATACGATGATGAAC 1799  
QY 735 CAAGAACGAATTAATGTCTCAGCCATCCGCAATGATGAGTGAAGTGGCGGCAACGTTG 794  
Db 1800 CAAGAACGAATGATGTTGCGACCATCCGCAATGATGAGTGAAGTGGCGGCAACGTTG 1859  
QY 795 CGTGGCATCGTCAATAACAGTTTGGAAACAACTCGAGGGCAGGCACTGCCACACATTT 854  
Db 1860 CGTGGCATCGTCAATAACAGTTTGGAAACAACTCGAGGGCAGGCACTGCCACACATTT 1919  
QY 855 CCAAAATGCCAATTCGGAGGAGCAGTACCGCAAGCGTTCTCGCTATTTCGCGCGGTGA 914  
Db 1920 CCAAAATGCCAATTCGGAGGAGCAGTACCGCAAGCGTTCTCGCTATTTCGCGCGGTGA 1979  
QY 915 TAAACAGACGAGGATTCGCGCTGATGCAACAGAGCGATTACGGCAACTTGTCTTACCA 974  
Db 1980 TAAACAGACGAGGATTCGCGCTGATGCAACAGAGCGATTACGGCAACTTGTCTTACCA 2039  
QY 975 CATCCGTAATAAACAATGCTTTTTCATTTTTCGGCAAGCAATGAGCGACAAGCTCAGCC 1034  
Db 2040 CATCCGTAATAAACAATGCTTTTTCATTTTTCGGCAAGCAATGAGCGACAAGCTCAGCC 2099







|||||  
Db 4260 CAACAGTAGCGCAACACAGCGGACGAGTCGGCGTAGGCTACCGGTTTC 4308

RESULT 8

AA543880  
ID AAS43880 standard; DNA; 3939 BP.  
XX  
AC AAS43880;  
XX  
DT 18-DEC-2001 (first entry)  
XX  
DE Neisseria meningitidis fusion protein delta-G741-983 DNA.  
XX  
KW Neisseria gonorrhoeae; leader peptide; fusion protein; ORF46.1; ds;  
KW Neisserial protein.  
XX  
OS Neisseria meningitidis.  
OS Synthetic.  
XX  
PN WO200164922-A2.  
XX  
PD 07-SEP-2001.  
XX  
PF 28-FEB-2001; 2001WO-IB00452.  
XX  
PR 28-FEB-2000; 2000GB-0004695.  
PR 13-NOV-2000; 2000GB-0027675.  
XX  
PA (CHIR-) CHIRON SPA.  
XX  
PI Arico MB, Comanducci M, Galeotti C, Maignani V, Guigliani MM;  
PI Pizza M;  
XX  
DR WPI; 2001-582163/65.  
DR P-PSDB; AAU27581.  
XX  
PT Producing heterologous proteins from Neisseria meningitidis and N.  
XX gonorrhoeae -  
XX  
PS Example 15; Page 51-52; 119pp; English.  
XX  
CC The invention relates to methods for the heterologous expression of  
CC Neisserial proteins from Neisseria meningitidis and Neisseria  
CC gonorrhoeae. At least one domain in the protein is deleted, e.g. the  
CC leader peptide, and may be replaced by a domain from a different protein  
CC to make a fusion protein, in order to enhance heterologous expression of  
CC Neisserial proteins. Also, a region of a protein, such as a poly-glycine  
CC stretch, can be mutated to enhance expression. The proteins used in the  
CC processes include ORF46.1, 287, 741, 919, 953, 961 and 983. Sequences  
CC AAS43868-AA543905 represent DNA molecules encoding Neisserial proteins  
CC and peptide regions of proteins of the invention.  
XX  
SQ Sequence 3939 BP; 1046 A; 1103 C; 1106 G; 684 T; 0 other;  
  
Query Match 93.1%; Score 2983.2; DB 22; Length 3939;  
Best Local Similarity 97.1%; Pred. No. 0;  
Matches 3064; Conservative 0; Mismatches 48; Indels 42; Gaps 1;  
  
QY 90 CGCGCGCGCGGCACCTCTCTGCGCCGACCTTCAATGCGAGCGCGCGGTATCGCGAGCAA 149  
|||||  
Db 759 CGCGCGAGCGCGCACCTCTGCGCCGACCTTCAATGCGAGCGGTACCGGTATCGCGAGCAA 818  
  
QY 150 CAGCAGAGCAACACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAAT 209  
|||||  
Db 819 CAGCAGAGCAACACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAAT 878  
  
QY 210 GTCAAAGACAGAGCATGCTCTGCGCGTGGGATGAGTTGCGGTACACACAGGGA 269  
|||||  
Db 879 GTCAAAGACAGAGCATGCTCTGCGCGTGGGATGAGTTGCGGTACACACAGGGA 938  
  
QY 270 TGCCAAATCAATGCCCGCCCGCCCGAATCTGCATACCGGAGCTTTTACAAACCCCAATGA 329  
|||||

Db 939 TGCCAAATCAATGCCCGCCCGCCCGAATCTGCATACCGGAGACTTTTCCAAACCCCAATGA 998  
QY 330 CGCATACAAGATTTTGCATCAACCTCAACCTGCAATTTGAAGCAGGCTATACAGACGCGG 389  
|||||  
Db 999 CGCATACAAGATTTTGCATCAACCTCAACCTGCAATTTGAAGCAGGCTATACAGACGCGG 1058  
QY 390 GGTAGAGTAGGTATCGTATACAGCGCAATCCGTCGGCAGCATATCCTTTTCCCGAACT 449  
|||||  
Db 1059 GGTAGAGTAGGTATCGTATACAGCGCAATCCGTCGGCAGCATATCCTTTTCCCGAACT 1118  
QY 450 GTATGCGAGAAAAGAACACAGCGCTATACGAAAATTTACAAAACATATACGCGGTATATGCG 509  
|||||  
Db 1119 GTATGCGAGAAAAGAACACAGCGCTATACGAAAATTTACAAAACATATACGCGGTATATGCG 1178  
QY 510 GAAGGAAGCGCTGAAGACGAGCGGTAAAGACATTTGAAGCTTCTTTCCGACGATGAGGC 569  
|||||  
Db 1179 GAAGGAAGCGCTGAAGACGAGCGGTAAAGACATTTGAAGCTTCTTTCCGACGATGAGGC 1238  
QY 570 CGTTATAGAGACTGAAGCAAGCCGACGGATATCCGCCACGTAAAGAAAATCGACACAT 629  
|||||  
Db 1239 CGTTATAGAGACTGAAGCAAGCCGACGGATATCCGCCACGTAAAGAAAATCGACACAT 1298  
QY 630 CGATGTGGTCTCCCATATATTTGGCGGGCTTCCGTGGAGCGGACGCTCGACGCGGTAT 689  
|||||  
Db 1299 CGATTTGGTCTCCCATATATTTGGCGGGCTTCCGTGGAGCGGACGCTCGACGCGGTAT 1358  
QY 690 TCGCGCCGATGCGACGCTACACATAATGATACGATGGAACCAAGAACGAAATAAT 749  
|||||  
Db 1359 TCGCGCCGATGCGACGCTACACATAATGATGGAACCAAGAACGAAATAATGAT 1418  
QY 750 GTCTGCGAGCATCCGCAATCATGGGTCAAGCTGGGCGAAACGTGGGTGCGCATCGTCAA 809  
|||||  
Db 1419 GGTGCGAGCATCCGCAATCATGGGTCAAGCTGGGCGAAACGTGGGTGCGCATCGTCAA 1478  
QY 810 TAACAGTTTTTGGAAACATCAGAGGCGGACGCTGCGGACCATTTTCCAAATAGCAATTC 869  
|||||  
Db 1479 TAACAGTTTTTGGAAACATCAGAGGCGGACGCTGCGGACCATTTTCCAAATAGCAATTC 1538  
QY 870 GGAGGAGCAGTACCGCAAGCGTTGCTCGCCTATTCGGCGGTGATAAAACACAGCAGGG 929  
|||||  
Db 1539 GGAGGAGCAGTACCGCAAGCGTTGCTCGCCTATTCGGCGGTGATAAAACACAGCAGGG 1598  
QY 930 TATCCGCTGATGCAACAGAGCGATTACGGCACTTGTCTTACACATCGTATATAAAA 989  
|||||  
Db 1599 TATCCGCTGATGCAACAGAGCGATTACGGCACTTGTCTTACACATCGTATATAAAA 1658  
QY 990 CATGCTTTTCAATTTTTCGCAAGCAATGACGACAAAGCTCAGCCCAACACACTGACCT 1049  
|||||  
Db 1659 CATGCTTTTCAATTTTTCGCAAGCAATGACGACAAAGCTCAGCCCAACACACTGACCT 1718  
QY 1050 ATTGCCATTTTATGAAAAGATGCTCAAAAGGCATTTATCAGATCGCAGCGGTAGACCG 1109  
|||||  
Db 1719 ATTGCCATTTTATGAAAAGATGCTCAAAAGGCATTTATCAGATCGCAGCGGTAGACCG 1778  
QY 1110 CAGTGGAGAAAAGTTCA-----A 1127  
|||||  
Db 1779 CAGTGGAGAAAAGTTCAAAACGGGAAATGATGGAGAACCGGTACAGACCGCTTGAGTA 1838  
QY 1128 TGGCTCAACCATTTGCGGAATTAATCTGCATGTTGTTGCTATCGGCACCCCTATGAAGCAAG 1187  
|||||  
Db 1839 TGGCTCAACCATTTGCGGAATTAATCTGCATGTTGTTGCTATCGGCACCCCTATGAAGCAAG 1898  
QY 1188 CGTCCGTTTCAACCGGTACAAACCGGATTCAAATTTCCCGGAACATCCTTTTCCCGACCCAT 1247  
|||||  
Db 1899 CGTCCGTTTCAACCGGTACAAACCGGATTCAAATTTCCCGGAACATCCTTTTCCCGACCCAT 1958  
QY 1248 CGTAAACCGGACGCGGCTCTGCTGCTGAGAAAATACCGTGGATGAGCAACGACCAACCT 1307  
|||||  
Db 1959 CGTAAACCGGACGCGGCTCTGCTGCTGAGAAAATACCGTGGATGAGCAACGACCAACCT 2018  
QY 1308 CGGTACACCGCTCTGCAACAGGCTCAGGACATCGGTGACGTGGGCGTGACAGCAAGTT 1367  
|||||  
Db 2019 CGGTACACCGCTCTGCAACAGGCTCAGGACATCGGTGCACTCGGCGTGACAGCAAGTT 2078

Qy	1368	CGGCTGGGACTGCTGGATCGGGTAAGGCCATGAACGACCCCGCTCTTTCGGTTCGG	1427
Db	2079	CGGCTGGGACGCTGCTGATCGGGTAAGGCCATGAACGACCCCGCTCTTTCGGTTCGG	2138
Qy	1428	CGACTTTACCGCGGATACGAAAGGTACATPCGATATGCCCTACTCTTCGTTCAGCAT	1487
Db	2139	CGACTTTACCGCGGATACGAAAGGTACATPCGATATGCCCTACTCTTCGTTCAGCAT	2198
Qy	1488	TTCAGCACGGGCGCTCATCAAAAAGGGCGAGCAACTGCAACTGCACGCGCAACAA	1547
Db	2199	TTCAGGCACGGGCGGCTGATCAAAAAGGGCGAGCAACTGCAACTGCACGCGCAACAA	2258
Qy	1548	CACCTATACGGGCAAAACCATTTATCGAAGGCGGTTCGCTGGTGTGTACGCGCAACAA	1607
Db	2259	CACCTATACGGGCAAAACCATTTATCGAAGGCGGTTCGCTGGTGTGTACGCGCAACAA	2318
Qy	1608	ATCGGATATGCGGTCGAAACCAAGGTGCGTGATTTATAACGGGGCGCATCCGGCGG	1667
Db	2319	ATCGGATATGCGGTCGAAACCAAGGTGCGTGATTTATAACGGGGCGCATCCGGCGG	2378
Qy	1668	TAGCCTGAACAGCGCGCATTTGCTATCTGCGACATACCACCGATCCGGCGCAACGA	1727
Db	2379	CAGCCTGAACAGCGCGCATTTGCTATCTGCGACATACCACCGAATCCGGCGCAACGA	2438
Qy	1728	AACCGTGCACATCAAAAGGCGATCTGCAGCTGGCGCGCAAGGTACGCTGTACACAGCTTT	1787
Db	2439	AACCGTGCACATCAAAAGGCGATCTGCAGCTGGCGCGCAAGGTACGCTGTACACAGCTTT	2498
Qy	1788	GGGCAAACTGCTGAAAGTGGAGGTACGGCGATACCGCGCGCAAGCTGTACATGTCCGC	1847
Db	2499	GGGCAAACTGCTGAAAGTGGAGGTACGGCGATTTATCGCGCGCAAGCTGTACATGTCCGC	2558
Qy	1848	ACGGGGCAAGGGCAGGCTATCTCAACGPTACCGGCACACGTTGCCCTTCGTAGTGC	1907
Db	2559	ACGGGGCAAGGGGCGAGCTATCTCAACAGTACCGCGCGAGCTGTTCCTTCCTTAGTGC	2618
Qy	1908	CGCCAAATCGGGCGGATTTATTTCTTCTTCAACAACATCGAAACCGACGGTGGTCTGCT	1967
Db	2619	CGCCAAATCGGGCGAGTATTTCTTCTTCAACAACATCGAAACCGACGGCGGCTGCT	2678
Qy	1968	GGCTTCCCTCGACAGCTCGAAAAACAGCGGCGAGTGAAGCGCACAGCTGTCTCTATTA	2027
Db	2679	GGCTTCCCTCGACAGCTCGAAAAACAGCGGCGAGTGAAGCGCACAGCTGTCTCTATTA	2738
Qy	2028	TGTCGCTCGGGCAATTCGGCGACGGACTGCTTCGGCAGGGGCACATTCGGCGCCGCGG	2087
Db	2739	TGTCGCTCGGGCAATTCGGCGACGGACTGCTTCGGCAGGGGCACATTCGGCGCCGCGG	2798
Qy	2088	TCTCAAAACGCGTAGACAGGGCGGACGAATCTGGAANAACCTGATGPTCGAACTTGA	2147
Db	2799	TCTGAANAACGCGTAGACAGGGCGGACGAATCTGGAANAACCTGATGPTCGAACTTGA	2858
Qy	2148	TGCTTCGGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCGCGCACACAGACA	2207
Db	2859	TGCTTCGGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCGCGCACACAGACA	2918
Qy	2208	TATCGCGGGCATCCGCCCTTACGGCGCAACTTTTCGGCGCAGGGCAGCCGTACAGCATGC	2267
Db	2919	TATCGCGGGCATCCGCCCTTACGGCGCAACTTTTCGGCGCAGGGCAGCCGTACAGCATGC	2978
Qy	2268	GAATGCCCGCGACGGTACGCATCTTCAACAGTCTCCCGCTACCGCTATGCCGACAG	2327
Db	2979	GAATGCCCGCGACGGTACGCATCTTCAACAGTCTCCCGCTACCGCTATGCCGACAG	3038
Qy	2328	TACCGCGCCCATGCCGATATGCAGGGACGCGGGTGAAGGCGCTATCGGACGGGTGGA	2387
Db	3039	TACCGCGCCCATGCCGATATGCAGGGACGCGCGCTGAAGGCGCTATCGGACGGGTGGA	3098
Qy	2388	CCACAACGCTACGGGTCGGCGCTATCGCGCAAAACCCCAACAGGACGGTGGAACTGGGA	2447
Db	3099	CCACAACGCGACGGTCTGCGCGCTATCGCGCAAAACCCCAACAGGACGGTGGAACTGGGA	3158

Qy	2448	ACAGGCGGTGTTGAAGGCAAAATGCGCGGCAAGTACCAAAACCGTCGCGCATTTGCCGGCAA	2507
Db	3159	ACAGGCGGTGTTGAAGGCAAAATGCGCGGCAAGTACCAAAACCGTCGCGCATTTGCCGGCAA	3218
Qy	2508	AACCGGCGAAATACGACAGCAGCCGCCACACTGGGCGATGGGACACGACATGAGCGGA	2567
Db	3219	AACCGGCGAAATACGACAGCAGCCGCCACACTGGGCGATGGGACACGACATGAGCGGA	3278
Qy	2568	AAACAGTGCAAAATGCAAAAACCCGACAGCATTTAGTCTGTTTGCAGGCATACGGCACGATGC	2627
Db	3279	AAACAGTGCAAAATGCAAAAACCCGACAGCATTTAGTCTGTTTGCAGGCATACGGCACGATGC	3338
Qy	2628	GGCGGATATCGGTATCTCAAAAGGCGCTGTTCTCTACGAGCGCTACAAAACACAGATCAG	2687
Db	3339	GGCGGATATCGGTATCTCAAAAGGCGCTGTTCTCTACGAGCGCTACAAAACACAGATCAG	3398
Qy	2688	CCGACGACCGGTGCGGACGGAACATCGGGAAGGACCGCTCAACGGCAGCGCTGATGCGAGCT	2747
Db	3399	CCGACGACCGGTGCGGACGGAACATCGGGAAGGACCGCTCAACGGCAGCGCTGATGCGAGCT	3458
Qy	2748	GGGCGCACTGGGCGGTGCTCAACGCTTCGCTTTCGCGCAACGGGAGATTTTGACGGTCTGAAGG	2807
Db	3459	GGGCGCACTGGGCGGTGCTCAACGCTTCGCTTTCGCGCAACGGGAGATTTTGACGGTCTGAAGG	3518
Qy	2808	CGGTCTCGGCTACGACCTGCTCAACAGAGTGCATTTCGCCGAAAAAGGCAGTGTCTTTGGG	2867
Db	3519	CGGTCTCGGCTACGACCTGCTCAACAGAGTGCATTTCGCCGAAAAAGGCAGTGTCTTTGGG	3578
Qy	2868	CTGAGGCGGCAACAGCCTCTACTGAAGGCACACTGTGTCGAGCTTCGGGGTCTGAAAGCTGTC	2927
Db	3579	CTGAGGCGGCAACAGCCTCTACTGAAGGCACACTGTGTCGAGCTTCGGGGTCTGAAAGCTGTC	3638
Qy	2928	GCAACCTTGAGCGATAAAGCCGCTCTGTTTGCACGGCGGCGTGGAAACGCGACCTGGA	2987
Db	3639	GCAACCTTGAGCGATAAAGCCGCTCTGTTTGCACGGCGGCGTGGAAACGCGACCTGGA	3698
Qy	2988	CGGACGGGACTTACACGGTAACGGGCGGCTTTTACCGGCGGACTGACGCAACCGCAAGAC	3047
Db	3699	CGGACGGGACTTACACGGTAACGGGCGGCTTTTACCGGCGGACTGACGCAACCGCAAGAC	3758
Qy	3048	GGGGCGACGCAATATCGGCGACACCCGCTGTTGCGCGTCTGGGCGGCGATGTCGAATT	3107
Db	3759	GGGGCGACGCAATATCGGCGACACCCGCTGTTGCGCGTCTGGGCGGCGATGTCGAATT	3818
Qy	3108	CGGCAACGGCTGGAACGGCTTGGCAGTTACAGCTACGCGGCTTCCAAACACAGTACGGCAA	3167
Db	3819	CGGCAACGGCTGGAACGGCTTGGCAGTTACAGCTACGCGGCTTCCAAACACAGTACGGCAA	3878
Qy	3168	CCACAGCGGACGAGTCGGCGTAGGCTACCGGTTTC	3201
Db	3879	CCACAGCGGACGAGTCGGCGTAGGCTACCGGTTTC	3912
RESULT	9		
AAID17045			
ID	AAID17045		
XX			
AC	AAID17045;		
XX			
DT	29-NOV-2001	(first entry)	
XX			
DE	N. meningitidis strain 2996	delta G741-983 fusion DNA.	
XX			
KW	Heterologous expression; Neisserial protein;		
KW	delta G741-983 fusion protein; ds.		
XX			
OS	Neisseria meningitidis 2996.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	1..3939	
FT		/*tag= a	
FT		/product= "N. meningitidis strain 2996	delta
FT		G741-983 fusion protein"	

XX	WO200164920-A2.	
PN		
XX		
PD	07-SEP-2001.	
XX		
XX	28-FEB-2001; 2001WO-IB00420.	
PF		
XX		
XX	28-FEB-2000; 2000GB-0004695.	
PR		
XX	13-NOV-2000; 2000GB-0027675.	
PR		
XX		
PA	(CHIR-) CHIRON SPA.	
XX		
PI	Arico MB, Comanducci M, Galeotti C, Masignani V, Giuliani MM;	
PI	Pizza M;	
XX		
DR	WPI: 2001-557776/62.	
DR	P-PSDB; AAE10030.	
XX		
XX		
PT	Heterologous expression for the expression of two or more Neisserial	
PT	proteins in fused state -	
XX		
XX	Example 4; Page 20-21; 52pp; English.	
XX		
CC	The present invention relates to a method for simultaneous heterologous	
CC	expression of two or more Neisserial proteins which are in a fused	
CC	state. The method is useful for simultaneous heterologous expression of	
CC	two or more Neisserial proteins. A protein that may be unstable or	
CC	poorly expressed on its own is assisted by adding a suitable hybrid	
CC	partner and commercial manufacture is simplified-only one expresseion and	
CC	purification need to be employed in order to produce two separately-	
CC	useful proteins. The present sequence is a DNA encoding	
CC	Neisseria meningitidis (serogroup B, strain 2996) delta G741-983	
CC	fusion protein.	
XX		
SQ	Sequence 3939 BP; 1046 A; 1103 C; 1106 G; 684 T; 0 other;	
Query Match 93.1%; Score 2983.2; DB 22; Length 3939;		
Best Local Similarity 97.1%; Pred. No. 0;		
Matches 3064; Conservative 0; Mismatches 48; Indels 42; Gaps 1;		
Qy	90 CGCGCGCGCGCGACTCTGCGCCCGACTTCAATGCAGCGCGCACCGGTATCGGCAGCAA	149
Db		
	759 CGCGCGAGCGCGCACTTCTGCGCCCGACTTCAATGCAGCGCGGTACCGGTATCGGCAGCAA	818
Qy	150 CAGCAGAGCAACAACAGCGGAATCAGCAGCAGTATCTTACGCCGGTATCAAGAAGCAAT	209
Db		
	819 CAGCAGACACACACGCGAATCAGCAGCAGTATCTTACGCCGGTATCAGAAGCAAT	878
Qy	210 GTGCAAGACAGACAGCATGCTCTGTGCGGTGCGGATGACGTGCGGTTCACAGACAGGA	269
Db	879 GTGCAAGACAGACAGCATGCTCTGTGCGGTGCGGATGACGTGCGGTTCACAGACAGGA	938
Qy	270 TGCCAAATCAATGCCCGCCCGCGAACTGCATACCGGAGACTTCAAAACCCAAATGA	329
Db		
	939 TGCCAAATCAATGCCCGCCCGCGAACTGCATACCGGAGACTTCAAAACCCAAATGA	998
Qy	330 CGCATACAAGAATTGTATCAACCTCAACCTGCAATTTGAAGCAGGCTATACAGACGCGG	389
Db		
	999 CGCATACAAGAATTGTATCAACCTCAACCTGCAATTTGAAGCAGGCTATACAGACGCGG	1058
Qy	390 GGTAGAGTAGGTATGTCGTGATACAGCGGATCCGTGCGCAGCATATCCTTTCCGGAAT	449
Db	1059 GGTAGAGTAGGTATGTCGTGATACAGCGGATCCGTGCGCAGCATATCCTTTCCGGAAT	1118
Qy	450 GTATGGCAGAAAAGACACGCGTATAACGAAAATACAAAACCTATACGGCGTATATCGG	509
Db		
	1119 GTATGGCAGAAAAGACACGCGTATAACGAAAATACAAAACCTATACGGCGTATATCGG	1178
Qy	510 GAAGGAAGCGCTGAAGACGGAGCGGTAAAGACATTAAGACTTCTTTTCGACGATGAGGC	569
Db	1179 GAAGGAAGCGCTGAAGACGGAGCGGTAAAGACATTAAGACTTCTTTTCGACGATGAGGC	1238
Qy	570 CGTTATAGAGACTGAAGCAAAAGCGGACGGATATCCGCCACGTAAAGAAATCGGACAT	629

1239	CGTTATAGACACTGAACAAAGCCGACGATATCGCCACCTGTAAGAAGAAATTCGACACAT	1298	
QY	CGATGTGCTCCCATATATTATTGGCGGGCGTTCCGTGGACGCGACACCTCGACGGCGTAT	689	
Db	CGATTTGGTCTCCCATATATTATTGGCGGGCGTTCCGTGGACGCGACACCTCGACGGCGTAT	1358	
QY	TGGCCCGGATGCGACGCTACACATATGAATACGATATGGAACCAAGACGAATATAT	749	
Db	TGGCCCGGATGCGACGCTACACATATGAATACGATATGGAACCAAGACGAATATAT	1418	
QY	GTCTGCAGCCATCCGCAATGCAATGGGTCAAGCTGGCGGAAACGTGGCGTGCGCATCGTCAA	809	
Db	GGTTGACAGCCATCCGAATGCAATGGGTCAAGCTGGCGGAAACGTGGCGTGCGCATCGTCAA	1478	
QY	TACAGTTTTTGGAAACAACATCGAGGCGAGCGACTGCCGACCATTTTCCAAATAGCCAATTC	869	
Db	TAACAGTTTTTGGAAACAACATCGAGGCGAGCGACTGCCGACCTTTTCCAAATAGCCAATTC	1538	
QY	GGAGGACGACTACCGCAAGCGTTGCTGCCTATTTCGGCGGTGATAAAACAGACGAGGG	929	
Db	GGAGGACGACTACCGCAAGCGTTGCTGCCTATTTCGGCGGTGATAAAACAGACGAGGG	1598	
QY	TATCCGCTGATGCAACAGAGCGATTACGGCAACTTTGCTTACACACATCGGTATATAAAA	989	
Db	TATCCGCTGATGCAACAGAGCGATTACGGCAACTTTGCTTACACACATCGGTATATAAAA	1658	
QY	CATCGTTTTCATTTTTTGGCAAGCAATCAGCGACAAGCTCAGCCCAACACACTGACCTT	1049	
Db	CATCGTTTTCATTTTTTGGCAAGCAATCAGCGACAAGCTCAGCCCAACACACTGACCTT	1718	
QY	ATTGCCATTTTATGAAAAGATGCTCAAAAAGGCAATTATCAGTCGCGAGCGTAGACCG	1109	
Db	ATTGCCATTTTATGAAAAGAGCTTCAAAAAGGCAATTATCAGTCGCGAGCGTAGACCG	1778	
QY	CAGTGGAGAAAAGTTCA -----A 1127		
Db	CAGTGGAGAAAAGTTCAAAACGGGAAATGATGGAGAACCGGGTACAGAACCGCTTGAGTA	1838	
QY	TGGCTCCAACCATTTGCGGAATTACTGCCATGTGTCCTATTCGGCACCTTATGAAGCAAG	1187	
Db	TGGCTCCAACCATTTGCGGAATTACTGCCATGTGTCCTATTCGGCACCTTATGAAGCAAG	1898	
QY	CGTCCGTTTCAACCGTACAAACCCGATTCAAATTGCCGGAACATCTTTTCCGCAACCAT	1247	
Db	CGTCCGTTTCAACCGTACAAACCCGATTCAAATTGCCGGAACATCTTTTCCGCAACCAT	1958	
QY	GGTAACCGGACGGCGCTCTGCTGTCGAGAAATACCCGTGGATGAGCAACGACACCT	1307	
Db	GGTAACCGGACGGCGCTCTGCTGTCGAGAAATACCCGTGGATGAGCAACGACACCT	2018	
QY	CGTACCACGCTGCTGACACGCTCAGGACATCGTGCAGTCGCGGTGGACAGCAAGTT	1367	
Db	CGTACCACGCTGCTGACACGCTCAGGACATCGGTGCAGTCGCGGTGGACAGCAAGTT	2078	
QY	CGCTGGGGACTGCTGATCGGGTAAGCCCATGAACGACCCGCTCTTTTCCGTTCCG	1427	
Db	CGCTGGGGACTGCTGATCGGGTAAGCCCATGAACGACCCGCTCTTTTCCGTTCCG	2138	
QY	CGACTTTACCGCGATACGAAAGGTACATCCGATATTTGCCCTACTCTTCCGTACGACAT	1487	
Db	CGACTTTACCGCGATACGAAAGGTACATCCGATATTTGCCCTACTCTTCCGTACGACAT	2198	
QY	TTACGACACGGCGGCTGATCAAAAAGCGGCGACCAACTGCAACTGCAAGCAACAA	1547	
Db	TTACGACACGGCGGCTGATCAAAAAGCGGCGACCAACTGCAACTGCAAGCAACAA	2258	
QY	CACCTATACGGGCAAAACCATTTATCGAAGCGGTTTCGCTGGTGTGTATACGGCAACAA	1607	
Db	CACCTATACGGGCAAAACCATTTATCGAAGCGGTTTCGCTGGTGTGTATACGGCAACAA	2318	
QY	ATCGGATATCGGCTCGAAACCAAGAGTGCCTGATTTATACGGGGGCGCATCCGCGG	1667	

Db 2319 ATCGGATATGCGCGTCGAAACCAAGGTCGCTGATTATTAACGGGGCGGCATCCGGCGG 2378  
Qy 1668 TAGCCTGAACGACGCGCATTTGCTATCTGGCAGATACCCAGCCATCGCGGCGCAACGA 1727  
Db 2379 CAGCCTGACGCGACGCGATGCTATCTGCGCATACCCAGCAATCCGGGCGAAACGA 2438  
Qy 1728 AACCGTGACATCAAGGCGATCTCAGCTGGGGCGGAAAGGTACGCTGTACACACGTTT 1787  
Db 2439 AACCGTACACATCAAGGCGAGTCTCAGCTGGAGCGCAAGGTACGCTGTACACACGTTT 2498  
Qy 1798 GGGCAAACTGCTGAAGTGGAGCGGTACGCGCATGACCGGCGCAAGCTGTACATGTCGC 1847  
Db 2499 GGGCAAACTGCTGAAGTGGAGCGGTACGCGCATGACCGGCGCAAGCTGTACATGTCGC 2558  
Qy 1848 ACGCGGCAAGGGGCGAGCTATCTCAACCGTACCGGCAACAGTGTTCCTCTCAGTGC 1907  
Db 2559 ACGCGGCAAGGGGCGAGCTATCTCAACAGTACCGGCAAGCTGTTCCTCTCAGTGC 2618  
Qy 1908 CGCCAAATCGGGCGGGATTATCTTCTTCAACAACTCGAAACCGAGGTGTCTGCT 1967  
Db 2619 CGCCAAATCGGGCGGGATTATCTTCTTCAACAACTCGAAACCGAGCGGGCGCTGCT 2678  
Qy 1968 GCTTCCCTCGACAGCTCGGAAAGAGGGGCGAGTGAAGCGGACAGCTGTCTATTA 2027  
Db 2679 GCTTCCCTCGACAGCTCGGAAAGAGGGGCGAGTGAAGCGGACAGCTGTCTATTA 2738  
Qy 2028 TGTCCGTTCGGGCAATCGGACGAGCTGCTTTCGCGAGCGGACATTCGCGCCGCGCGG 2087  
Db 2739 TGTCCGTTCGGGCAATCGGACGAGCTGCTTTCGCGAGCGGACATTCGCGCCGCGCGG 2798  
Qy 2088 TGTGAACACGCGGTAGAACAGGGCGGAGCAATCTGGAACACCTGATGGTGAAGTGA 2147  
Db 2799 TGTGAACACGCGGTAGAACAGGGCGGAGCAATCTGGAACACCTGATGGTGAAGTGA 2858  
Qy 2148 TGCCTCCGAATCATCCGCAACACCGGACGAGCTTGAAGTTCGGGCGCGGACGACAGA 2207  
Db 2859 TGCCTCCGAATCATCCGCAACACCGGAGAGCTTGAAGTTCGGGCGCGGACGACAGA 2918  
Qy 2208 TATGCGGGGATCCGCCCTACGGGCGAACTTTCGGCGAGGCGGAGCGTACAGCATGC 2267  
Db 2919 TATGCGGGGATCCGCCCTACGGGCGAACTTTCGGCGAGGCGGAGCGGTACAGCATGC 2978  
Qy 2268 GAATGCCGCGAGCGGTACGCAATCTCAACAGTCTCCCGCTACCGTCTATGCGGACAG 2327  
Db 2979 GAATGCCGCGAGCGGTACGCAATCTCAACAGTCTCCCGCTACCGTCTATGCGGACAG 3038  
Qy 2328 TACCGCGCCCATGCGGATATGACGAGGAGCGGCGGTGAAGCCGTATCGAGCGGTTGA 2387  
Db 3039 TACCGCGCCCATGCGGATATGACGAGGAGCGGCGGTGAAGCCGTATCGAGCGGTTGA 3098  
Qy 2388 CCACAAACGCTACGGGTCTCGCGTCTATCGCGCAACCCAGAGCGGTGAAGCGTGGGA 2447  
Db 3099 CCACAAACGCGAGGTCTGCGGCTATCGCGCAACCCAGAGCGGTGAAGCGTGGGA 3158  
Qy 2448 ACAGGCGGTGTTGAAGCAAAATCGCGGCGAGTACCCAAACCGTCGGCATTCGCGCAA 2507  
Db 3159 ACAGGCGGTGTTGAAGCAAAATCGCGGCGAGTACCCAAACCGTCGGCATTCGCGCAA 3218  
Qy 2508 RACCGGCAAAATACGACAGCGCGCCACACTGGGCGATGGGACACAGCATGAGGAGA 2567  
Db 3219 RACCGGCAAAATACGACAGCGCGCCACACTGGGCGATGGGACACAGCATGAGGAGA 3278  
Qy 2568 AAACAGTCAAAATGCAAAACCGCAGCATTAGTCTGTTGAGGCGATACGCGAGTGC 2627  
Db 3279 AAACAGTCAAAATGCAAAACCGCAGCATTAGTCTGTTGAGGCGATACGCGAGTGC 3338  
Qy 2628 GGGCGATATCGGCTATCTCAAGGCGCTTCTCTTACGAGCGCTACAAAACAGCATCAG 2687  
Db 3339 GGGCGATATCGGCTATCTCAAGGCGCTTCTCTTACGAGCGCTACAAAACAGCATCAG 3398  
Qy 2688 CCGCAGCAGCGGTTCGGGACGACATGCGGAGGCGAGCTCAACGCGAGCTGATCGAGCT 2747  
Db 3399 CCGCAGCAGCGGTTCGGGACGACATGCGGAGGCGAGCTCAACGCGAGCTGATCGAGCT 3458

Qy 2748 GGGCGACTGGGCGGTGTCAACGTTTCCTTTCCCGCAACGGGAGATTTACGCTGAAGG 2807  
Db 3459 GGGCGACTGGGCGGTGTCAACGTTTCCTTTCCCGCAACGGGAGATTTACGCTGAAGG 3518  
Qy 2808 CGGTCTGCGCTACGACCTGCTCAAAACAGATGATTCGCCGAAAAAGGACGTCCTTGGG 2867  
Db 3519 CGGTCTGCGCTACGACCTGCTCAAAACAGATGATTCGCCGAAAAAGGACGTCCTTGGG 3578  
Qy 2868 CTGGAGCGGCAACGCTCACTGAAGGACACACTGTCGAGCTCGCGGCTCTGAAGCTGTC 2927  
Db 3579 CTGGAGCGGCAACGCTCACTGAAGGACGCTGTCGAGCTCGCGGCTCTGAAGCTGTC 3638  
Qy 2928 GCAACCCCTTGAGCGGATAAAGCGCTCTGTTGCAACGGCGGCGTGAACGCACTGAA 2987  
Db 3639 GCAACCCCTTGAGCGGATAAAGCGCTCTGTTGCAACGGCGGCGTGAACGCACTGAA 3698  
Qy 2988 CGGACGCGACTACAGGTAAACGGGCGGCTTTACCGCGCGGCTGACAGCAACGGCAAGAC 3047  
Db 3699 CGGACGCGACTACAGGTAAACGGGCGGCTTTACCGCGCGGCTGACAGCAACGGCAAGAC 3758  
Qy 3048 GGGGCGACGCAATATGCGCGCACACCCGCTGCTGTCGCGGCTGCGGCGGATGTCGAATT 3107  
Db 3759 GGGGCGACGCAATATGCGCGCACACCCGCTGCTGTCGCGGCTGCGGCGGATGTCGAATT 3818  
Qy 3108 CGGCAACCGCTGGAACGCTTTGGCAGCTTACAGCTACGCGCGGTTCCAAACAGTACGGCAA 3167  
Db 3819 CGGCAACCGCTGGAACGCTTTGGCAGCTTGGCAGCTTACAGCTACGCGCGTCCAAACAGTACGGCAA 3878  
Qy 3168 CCACAGCGGAGTGCAGGCGGTAGGCTACCGGTTTC 3201  
Db 3879 CCACAGCGGAGTGCAGGCGGTAGGCTACCGGTTTC 3912

## RESULT 10

AAS43902

ID AAS43902 standard; DNA; 4170 BP.

XX AC AAS43902;

XX DT 18-DEC-2001 (first entry)

XX DE Neisseria meningitidis fusion protein 961c-983 DNA.

XX DE Neisseria gonorrhoeae; leader peptide; fusion protein; ORF46.1; ds;

XX KW Neisserial protein.

XX OS Neisseria meningitidis.

XX OS Synthetic.

XX PN W0200164922-A2.

XX PD 07-SEP-2001.

XX XX 28-FEB-2001; 2001WO-IB00452.

XX XX 28-FEB-2000; 2000GB-0004695.

XX PR 13-NOV-2000; 2000GB-0027675.

XX XX (CHIR-) CHIRON SPA.

XX PI Arico MB, Comanducci M, Galeotti C, Massignani V, Guilianini MM;

XX PI Pizza M;

XX XX WPI; 2001-582163/65.

XX DR P-PSDB; AAU27605.

XX PT Producing heterologous proteins from Neisseria meningitidis and N.

XX PT gonorrhoeae -

XX XX Example 23; Page 71-73; 119pp; English.

XX CC The invention relates to methods for the heterologous expression of

CC Neisserial proteins from Neisseria meningitidis and Neisseria gonorrhoeae. At least one domain in the protein is deleted, e.g. the CC leader peptide, and may be replaced by a domain from a different protein CC to make a fusion protein, in order to enhance heterologous expression of CC Neisserial proteins. Also, a region of a protein, such as a poly-glycine CC stretch, can be mutated to enhance expression. The proteins used in the CC processes include ORF46.1, 287, 741, 919, 961 and 983. Sequences CC AAS43868-AAS43905 represent DNA molecules encoding Neisserial proteins CC and peptide regions of the invention.

XX  
SQ Sequence 4170 BP; 1185 A; 1138 C; 1119 G; 728 T; 0 other;

Query Match 93.1%; Score 2983.2; DB 22; Length 4170;  
Best Local Similarity 97.1%; Pred. No. 0;  
Matches 3064; Conservative 0; Mismatches 48; Indels 42; Gaps 1;

QY 90 CGCGCGCGCGGCACATCTTCGCCGCCACATTCATGTCAGCGCGGCACCGGTATCGGCAGCAA 149  
DB 990 CGCGCGAGCGCGACATCTTCGCCGCCACATTCATGTCAGCGCGGTATCGGCAGCAA 1049  
QY 150 CAGCAGAGCAACAACAGCGGAAATCAGCAGCATCTTACGCCGGGTATCAAGAACGAAAT 209  
DB 1050 CAGCAGAGCAACAACAGCGGAAATCAGCAGCATCTTACGCCGGGTATCAAGAACGAAAT 1109  
QY 210 GTGCAAGAGCAGAGCATGCTCTGTCCGGTCCGGATGACGTTGCGGTTACAGACAGGGA 269  
DB 1110 GTGCAAGAGCAGAGCATGCTCTGTCCGGTCCGGATGACGTTGCGGTTACAGACAGGGA 1169  
QY 270 TGCAGAAATCAATGCCCCCCCCCGAATCTGCATACCGGAGACTTTACAAACCAATGA 329  
DB 1170 TGCAGAAATCAATGCCCCCCCCCGAATCTGCATACCGGAGACTTTTCAAAACCAATGA 1229  
QY 330 CGCATACAAGAAATTTGATCAACCTCAAACTGCAATTTGAAGCAGGCTATACAGGACGCGG 389  
DB 1230 CGCATACAAGAAATTTGATCAACCTCAAACTGCAATTTGAAGCAGGCTATACAGGACGCGG 1289  
QY 390 GGTAGAGTAGGTATCGTTCGATACAGGCGAATCCGTCGCAGCATATCTTTCCCGAACT 449  
DB 1290 GGTAGAGTAGGTATCGTTCGATACAGGCGAATCCGTCGCAGCATATCTTTCCCGAACT 1349  
QY 450 GTATGCCAGAAAGACACGGCTATACGAAATTAACGAAATTAACGAAATTAACGAAATTAACG 509  
DB 1350 GTATGCCAGAAAGACACGGCTATACGAAATTAACGAAATTAACGAAATTAACGAAATTAACG 1409  
QY 510 GAAGGAAGCGCCTGAAGACGGAGCGGTAAAGACATTTAAAGCTTCTTTTCAGCATCAGGC 569  
DB 1410 GAAGGAAGCGCCTGAAGACGGAGCGGTAAAGACATTTAAAGCTTCTTTTCAGCATCAGGC 1469  
QY 570 CGTTATAGAGACTGAAGCAAGCGCAGGATATCCGCCACGTAAAGAAATCGGACACAT 629  
DB 1470 CGTTATAGAGACTGAAGCAAGCGCAGGATATCCGCCACGTAAAGAAATCGGACACAT 1529  
QY 630 CGATGTGGTCTCCCATATTTATTTGGCGGGGCTTCCTGTGGACGCGAGACCTTGCAGGCGGTAT 689  
DB 1530 CGATTTGGTCTCCCATATTTATTTGGCGGGGCTTCCTGTGGACGCGAGACCTTGCAGGCGGTAT 1589  
QY 690 TCGCGCCGATGCGACGCTACACATATGAATACGATGATGAACCAAGCAAGCAAAATAAT 749  
DB 1590 TCGCGCCGATGCGACGCTACACATATGAATACGATGATGAACCAAGCAAGCAAAATAAT 1649  
QY 750 GTCTGCAAGCCATCCGCAATGATGGGTCAAGCTTGGCGCAACGTGGCGTCGCGCATCGTCAA 809  
DB 1650 GTTTGCAAGCCATCCGCAATGATGGGTCAAGCTTGGCGCAACGTGGCGTCGCGCATCGTCAA 1709  
QY 810 TAACAGTTTTTGAACACATCAGGCGCAGCATGCGGACCATTTTCCAAATAGCCAAATTC 869  
DB 1710 TAACAGTTTTTGAACACATCAGGCGCAGCATGCGGACCATTTTCCAAATAGCCAAATTC 1769  
QY 870 GGAGGAGCAGTAGTCCGCAAGCGTTCGTCGCTATTTCCGCGCGGTGATTAACACAGCAGGG 929  
DB 1770 GGAGGAGCAGTAGTCCGCAAGCGTTCGTCGCTATTTCCGCGCGGTGATTAACACAGCAGGG 1829  
QY 930 TATCGCGCTGATGCAACAGAGCGGATTAACGGCAACTTGTCTTACCACATCCGTAATAAAAA 989

DB 1830 TATCGCGCTGATGCAACAGAGCGATTAACGCAACCTGTCTTACCACATCGTATATAAAA 1889  
QY 990 CATGCTTTTCATTTTTTTCGCAAGCAATGACGCAAGCTTACGCCCAACACACTGACCCCT 1049  
DB 1890 CATGCTTTTCATTTTTTTCGCAAGCAATGACGCAAGCTTACGCCCAACACACTGACCCCT 1949  
QY 1050 ATTGCCATTTTATGAAAAAGATGCTCAAAAAGCATTTATCAGATGCGAGCGGTAGACCG 1109  
DB 1950 ATTGCCATTTTATGAAAAAGATGCTCAAAAAGCATTTATCAGATGCGAGCGGTAGACCG 2009  
QY 1110 CAGTGGAGAAAAAGTTCA-----A 1127  
DB 2010 CAGTGGAGAAAAAGTTCAAAACGGGAAATGTATGGAGAACCGGTTACAGAACCGCTTGAGTA 2069  
QY 1128 TGCGCTCAACCATTTGGGAATTAATGCGCATGTGGTGCCTATTCGGCACCCCTATGAAGCAAG 1187  
DB 2070 TGCGCTCAACCATTTGGGAATTAATGCGCATGTGGTGCCTGTGCGCACCCCTATGAAGCAAG 2129  
QY 1188 CGTCCGTTTACCCGTCACAAACCCGATTCAAATTTGCCGGAACATCTCTTTTCCGACCCAT 1247  
DB 2130 CGTCCGTTTACCCGTCACAAACCCGATTCAAATTTGCCGGAACATCTCTTTTCCGACCCAT 2189  
QY 1248 CGTAACCGCGACCGCGGCTCTGCTGCTGCAGAAATACCCGTTGGATGAGCAACGACACCT 1307  
DB 2190 CGTAACCGCGACCGCGGCTCTGCTGCTGCAGAAATACCCGTTGGATGAGCAACGACACCT 2249  
QY 1308 GCGTACACGCTCTCTCACAACCGCTCAGGACATCGGTGCGAGTCGGGTGGACAGCAAGTT 1367  
DB 2250 GCGTACACGCTCTCTCACAACCGCTCAGGACATCGGTGCGAGTCGGGTGGACAGCAAGTT 2309  
QY 1368 CGGCTGGGACTCTCTGATGCGGGTAAAGCCATGAACGACCGCGTCTCTTCCGTTCCG 1427  
DB 2310 CGGCTGGGACTCTCTGATGCGGGTAAAGCCATGAACGACCGCGTCTCTTCCGTTCCG 2369  
QY 1428 CGACTTTTACCGCGATACGAAAGGTACATCCGATATTCCTTCTCTCTTCCGTAACGACAT 1487  
DB 2370 CGACTTTTACCGCGATACGAAAGGTACATCCGATATTCCTTCTCTCTTCCGTAACGACAT 2429  
QY 1488 TTCAGCAGCGGGCGGCTGATCAAAAAGCGCGCAACTGCAACTGCACGGCAACAA 1547  
DB 2430 TTCAGCAGCGGGCGGCTGATCAAAAAGCGCGCAACTGCAACTGCACGGCAACAA 2489  
QY 1548 CACTATACGGGCAAAACCATTTATCGAAGCGGTTTCGCTGCTGTGTGTACGGCAACAA 1607  
DB 2490 CACTATACGGGCAAAACCATTTATCGAAGCGGTTTCGCTGCTGTGTGTACGGCAACAA 2549  
QY 1608 ATCGGATATCGCGGTGCAAAACCAAGGTGCGCTGATTTTATAACGGGGCGGSCATCCGCGG 1667  
DB 2550 ATCGGATATCGCGGTGCAAAACCAAGGTGCGCTGATTTTATAACGGGGCGGSCATCCGCGG 2609  
QY 1668 TAGCCTGAACAGCGCGGATTTGCTTATCTGCGAGATACGGACCGATCCGGCGCAACGA 1727  
DB 2610 TAGCCTGAACAGCGCGGATTTGCTTATCTGCGAGATACGGACCGATCCGGCGCAACGA 2669  
QY 1728 AACCGTGCACATCAAAAGCGATCTGAGCTGGCGGCGAAGGTACGCTGTACACAGCTTT 1787  
DB 2670 AACCGTGCACATCAAAAGCGATCTGAGCTGGAGCGCAAGGTACGCTGTACACAGCTTT 2729  
QY 1788 GGCAAACTGCTGAAAGTGGACGCTGAGCGGATGACGGCGGCAAGCTGTACATGTGCGG 1847  
DB 2730 GGCAAACTGCTGAAAGTGGACGCTGAGCGGATGACGGCGGCAAGCTGTACATGTGCGG 2789  
QY 1848 ACAGCGCAAGGGCAGGCTATCTCAACCGTACCGGACACAGCTTTCCTTCTGAGTGC 1907  
DB 2790 ACAGCGCAAGGGCAGGCTATCTCAACCGTACCGGACACAGCTTTCCTTCTGAGTGC 2849  
QY 1908 CGCAAAATCGGCGGATTTCTTTCTCAAAACATCGAAACCGAGCGGTGCTGCT 1967  
DB 2850 CGCAAAATCGGCGGATTTCTTTCTCAAAACATCGAAACCGAGCGGCTGCT 2909  
QY 1968 GGTCTCCCTCGACAGCGTCAAAAAACAGCGGCACTGAAGGCGACACGCTGTCTTATTTA 2027

Db 2910 GGCTTCCTCGACAGCGTGAAGGCGAGCGAGTGAAGGCGAGCGTGTCTCTATTA 2969  
Qy 2028 TGTCCGTGCGGCAATCGGCACGAGCTGCTTGGCAGCGGACATTCGGCGCGCGCGG 2087  
Db 2970 TGTCCGTGCGGCAATCGGCACGAGCTGCTTGGCAGCGGACATTCGGCGCGCGCGG 3029  
Qy 2088 TCTGAACACGCGGTAGAACAGGCGGAGCAATCTGGAACACCTGTAGTGGTCGAACCTGGA 2147  
Db 3030 TCTGAACACGCGGTAGAACAGGCGGAGCAATCTGGAACACCTGTAGTGGTCGAACCTGGA 3089  
Qy 2148 TGCCTCCGAATCATCCGCAACACCGGAGAGCGTTGAAACTTGGGCGCGCGAGCCGACAGCA 2207  
Db 3090 TGCCTCCGAATCATCCGCAACACCGGAGAGCGTTGAAACTTGGGCGCGCGAGCCGACAGCA 3149  
Qy 2208 TATGCCGGGCAATCCGCGCTAGCGGCAACTTTCGCGCAGCGGAGCGGAGCGGATACGATGC 2267  
Db 3150 TATGCCGGGCAATCCGCGCTAGCGGCAACTTTCGCGCAGCGGAGCGGAGCGGATACGATGC 3209  
Qy 2268 GAATGCCCGGACGCTACGATCTTCAACAGTCTCGCGCTACCGTCTATGCGGACAG 2327  
Db 3210 GAATGCCCGGACGCTACGATCTTCAACAGTCTCGCGCTACCGTCTATGCGGACAG 3269  
Qy 2328 TACCGCGCGCATGCCGATATGACGAGGAGCGGCTGAAAGCCGTATCGGAGGCTTGA 2387  
Db 3270 TACCGCGCGCATGCCGATATGACGAGGAGCGGCTGAAAGCCGTATCGGAGGCTTGA 3329  
Qy 2388 CCACAAACGCTACGGGTCTCGCGTCTATCGCGCAACCAACAGGAGCGTGAACGTGGGA 2447  
Db 3330 CCACAAACGCTACGGGTCTCGCGTCTATCGCGCAACCAACAGGAGCGTGAACGTGGGA 3389  
Qy 2448 ACAGGCGGCTGTGAAGGCAAAATCGCGGCGAGTACCCAAACCCGTCGGCATTTGCCGCGAA 2507  
Db 3390 ACAGGCGGCTGTGAAGGCAAAATCGCGGCGAGTACCCAAACCCGTCGGCATTTGCCGCGAA 3449  
Qy 2508 AACCGGCAAAATACGACAGAGCGCGCACACTGGGATGGGACACAGCATGAGGGA 2567  
Db 3450 AACCGGCAAAATACGACAGAGCGCGCACACTGGGATGGGACACAGCATGAGGGA 3509  
Qy 2568 AAACAGTCAAAATGCAAAACCGGAGGATAGTCTGTTGAGGATACGAGCATGAGGATGC 2627  
Db 3510 AAACAGTCAAAATGCAAAACCGGAGGATAGTCTGTTGAGGATACGAGCATGAGGATGC 3569  
Qy 2628 GGGGATATCGGCTATCTCAAGGCGCTTCTCTACGAGCGCTACAAACACAGCATCAG 2687  
Db 3570 GGGGATATCGGCTATCTCAAGGCGCTTCTCTACGAGCGCTACAAACACAGCATCAG 3629  
Qy 2688 CGGAGACCGGTGCGGAGAAATCGGAGGAGCGGATCAACGCGAGCTGATGACGCT 2747  
Db 3630 CGGAGACCGGTGCGGAGAAATCGGAGGAGCGGATCAACGCGAGCTGATGACGCT 3689  
Qy 2748 GGGGCACTGGGCGGTGTCACGCTTCCGTTTCCGCAACGGGAGATTTGACGGTCAAGG 2807  
Db 3690 GGGGCACTGGGCGGTGTCACGCTTCCGTTTCCGCAACGGGAGATTTGACGGTCAAGG 3749  
Qy 2808 CGGTGCGCTACGACCTGCTCAAAAGAGATGATTCGCCGCAAAAGGAGCTGCTTTGGG 2867  
Db 3750 CGGTGCGCTACGACCTGCTCAAAAGAGATGATTCGCCGCAAAAGGAGCTGCTTTGGG 3809  
Qy 2868 CTGGAGCGCAACGCTCACTGAAGGACACACTGTGCGAGCTCGGGTCTGAACCTGTC 2927  
Db 3810 CTGGAGCGCAACGCTCACTGAAGGACACACTGTGCGAGCTCGGGTCTGAACCTGTC 3869  
Qy 2928 GCAACCCCTTGAGCGATAAGCGCTCTGTTTGAACGCGGCGGTGGAACGCACTGAA 2987  
Db 3870 GCAACCCCTTGAGCGATAAGCGCTCTGTTTGAACGCGGCGGTGGAACGCACTGAA 3929  
Qy 2988 CGGACGCGACTACAGGTAACGGGCGGCTTTTACCGGCGCGACTGCGAGCAACCGGCAAGAC 3047  
Db 3930 CGGACGCGACTACAGGTAACGGGCGGCTTTTACCGGCGCGACTGCGAGCAACCGGCAAGAC 3989  
Qy 3048 GGGGCGACGAATATGCGGACACACCGGCTGTTGCGCGGTCTGGGCGGCGGATGTCGAATT 3107  
Db 3990 GGGGCGACGAATATGCGGACACACCGGCTGTTGCGCGGTCTGGGCGGCGGATGTCGAATT 4049

Qy 3108 CGGCAACGCGCTGGAACGCGTGGACAGTTTACAGCTTACGCCGCTTCCAACACGATACGGCAA 3167  
Db 4050 CGGCAACGCGCTGGAACGCGTGGACAGTTTACAGCTTACGCCGCTTCCAACACGATACGGCAA 4109  
Qy 3168 CCACAGCGGACGAGTCGCGGTAGCTACCGGTTTC 3201  
Db 4110 CCACAGCGGACGAGTCGCGGTAGCTACCGGTTTC 4143

RESULT 11  
AAD17055  
ID AAD17055 standard; DNA; 4170 BP.  
XX AAD17055;  
AC AAD17055;  
XX  
DT 29-NOV-2001 (first entry)  
XX  
DE N. meningitidis strain 2996 961c-983 fusion DNA.  
XX  
KW Heterologous expression; Neisserial protein;  
delta 961c-983 fusion protein; ds.  
XX  
OS Neisseria meningitidis 2996.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..4170  
FT /\*tag= a  
FT /product= "N. meningitidis strain 2996 961c-983  
fusion protein"  
XX  
PN WO200164920-A2.  
XX  
PD 07-SEP-2001.  
XX  
XX 28-FEB-2001; 2001WO-IB00420.  
XX  
XX 28-FEB-2000; 2000GB-0004695.  
PR 13-NOV-2000; 2000GB-0027675.  
XX  
XX (CHIR-) CHIRON SPA.  
XX

PI Arico MB, Comanducci M, Galeotti C, Masignani V, Giuliani MM;  
PI Pizza M;  
XX  
XX WPI: 2001-557776/62.  
DR P-PSDB; AAE10040.  
XX  
PT Heterologous expression for the expression of two or more Neisserial  
proteins in fused state  
XX  
PS Example 23; Page 31; 52pp; English.  
XX  
CC The present invention relates to a method for simultaneous heterologous  
expression of two or more Neisserial proteins which are in a fused  
state. The method is useful for simultaneous heterologous expression of  
two or more Neisserial proteins. A protein that may be unstable or  
poorly expressed on its own is assisted by adding a suitable hybrid  
partner and commercial manufacture is simplified-only one expression and  
purification need to be employed in order to produce two separately-  
useful proteins. The present sequence is a DNA encoding  
CC Neisseria meningitidis (serogroup B, strain 2996) 961c-983 fusion  
protein.  
XX  
SQ Sequence 4170 BP; 1185 A; 1138 C; 1119 G; 728 T; 0 other;

Query Match 93.1%; Score 2983.2; DB 22; Length 4170;  
Best Local Similarity 97.1%; Pred. No. 0;  
Matches 3064; Conservative 0; Mismatches 48; Indels 42; Gaps 1;  
Qy 90 CGCGGCGGCGGCGACTTCTGGCGCGGAGCTTCAATGCGGCGGCGGATCGGATCGGAGCAA 149  
Db 990 CGCGGCGGCGGCGACTTCTGGCGCGGAGCTTCAATGCGGCGGCGGATCGGATCGGAGCAA 1049





```
|||||
Db 3210 GAATCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCGCACAG 3269
QY 2328 TACCCGCCCATCGGATATTCACGGGACGCGGCTGAAGCCGTATCGGCGGTTTGA 2387
Db 3270 TACCCGCCCATCGGATATTCACGGGACGCGGCTGAAGCCGTATCGGCGGTTTGA 3329
QY 2388 CCACAAACGCTACGGCTCTCGCGCTCATCGCGCAAAACCCAAACAGGACGGTGAACCTGGGA 2447
Db 3330 CCACAAACGCGAGGCTCTCGCGCTCATCGCGCAAAACCCAAACAGGACGGTGAACCTGGGA 3389
QY 2448 ACAGGGCGGTGTTGAAGGCAAAATCGCGGCGAGTACCCAAACCCGTCGGCATTTGCCGCGAA 2507
Db 3390 ACAGGGCGGTGTTGAAGGCAAAATCGCGGCGAGTACCCAAACCCGTCGGCATTTGCCGCGAA 3449
QY 2508 AACCGCGCAAAATACGACGAGCGCGCACACTGGGCGATGGGACACACACACATGAGCGA 2567
Db 3450 AACCGCGCAAAATACGACGAGCGCGCACACTGGGCGATGGGACACACACATGAGCGA 3509
QY 2568 AAACAGTCAATGCAAAACCCGACGAGCATTTAGTCTGTTGACGACATACGGCAGCATGC 2627
Db 3510 AAACAGTCAATGCAAAACCCGACGAGCATTTAGTCTGTTGACGACATACGGCAGCATGC 3569
QY 2628 GGGCGATATCGGCTATCTCAAGAGCGCTGTCTCTACGAGCGCTACAAAACAGCATCAG 2687
Db 3570 GGGCGATATCGGCTATCTCAAGAGCGCTGTCTCTACGAGCGCTACAAAACAGCATCAG 3629
QY 2688 CCGCAGCACCCTGCGGACGACATGCGGAAGGCGAGCTCAACGCGACGCTGATCGACCT 2747
Db 3630 CCGCAGCACCCTGCGGACGACATGCGGAAGGCGAGCTCAACGCGACGCTGATCGACCT 3689
QY 2748 GGGCGCACTGGCGGTGTCAACGCTTCCGTTTCCGCAACGCGGAGATTTGACGGTCAAGG 2807
Db 3690 GGGCGCACTGGCGGTGTCAACGCTTCCGTTTCCGCAACGCGGAGATTTGACGGTCAAGG 3749
QY 2808 CGGTCTGCGCTACGACCTGCTCTCAACAGAGATGATTCGCCGCAAAAGGCGCTTTGGG 2867
Db 3750 CGGTCTGCGCTACGACCTGCTCTCAACAGAGATGATTCGCCGCAAAAGGCGCTTTGGG 3809
QY 2868 CTGGAGCGGCAACGCTCACTGAAGGCACACTGCTCGGACCTGCGGGTCTGAAGCTGTC 2927
Db 3810 CTGGAGCGGCAACGCTCACTGAAGGCACACTGCTCGGACCTGCGGGTCTGAAGCTGTC 3869
QY 2928 GCAACCCCTTGACGATAAAGCGCTCTGTTTGAACGGCGGCGTGAACGCGACCTGAA 2987
Db 3870 GCAACCCCTTGACGATAAAGCGCTCTGTTTGAACGGCGGCGTGAACGCGACCTGAA 3929
QY 2988 CGGACGCGACTACACGGTAAACGGCGGCTTTTACCGGGCGGACTGCAAGCAACCGGCAAGC 3047
Db 3930 CGGACGCGACTACACGGTAAACGGCGGCTTTTACCGGGCGGACTGCAAGCAACCGGCAAGC 3989
QY 3048 GGGGCGACCAATATGCCGCACACCCGCTGGTTCGCGGCTGTCGGCGGAGTGTGCAATT 3107
Db 3990 GGGGCGACCAATATGCCGCACACCCGCTGGTTCGCGGCTGTCGGCGGAGTGTGCAATT 4049
QY 3108 CGGCAACGCTGGAACGGCTTGGCAGTTACAGCTACGCGCGGTTCCAAACAGTACGGCAA 3167
Db 4050 CGGCAACGCTGGAACGGCTTGGCAGTTACAGCTACGCGCGGTTCCAAACAGTACGGCAA 4109
QY 3168 CCACAGGCGAGTTCGGCGTGTAGGTACCGGTTTC 3201
Db 4110 CCACAGGCGAGTTCGGCGTGTAGGTACCGGTTTC 4143
```

RESULT 12  
AAS43875  
ID AAS43875 standard; DNA; 3939 BP.  
XX  
AC AAS43875;  
XX  
DT 18-DEC-2001 (first entry)  
XX  
DE Neisseria meningitidis fusion protein delta-G983-741 DNA.

```
xx Neisseria gonorrhoeae; leader peptide; fusion protein; ORF46.1; ds;  
kw Neisserial protein.  
xx  
os Neisseria meningitidis.  
os Synthetic.  
xx  
pn WO200164922-A2.  
xx  
pd 07-SEP-2001.  
xx  
pf 28-FEB-2001; 2001WO-IB00452.  
xx  
pr 28-FEB-2000; 2000GB-0004695.  
pr 13-NOV-2000; 2000GB-0027675.  
xx  
pa (CHIR-) CHIRON SPA.  
xx  
pi Arico MB, Comanducci M, Galeotti C, Massignani V, Guillian MM;  
pi Pizzo M;  
xx  
dr WPI: 2001-582163/65.  
dr P-FSDB; AAU27576.  
xx  
pt Producing heterologous proteins from Neisseria meningitidis and N.  
pt gonorrhoeae -  
xx  
ps Example 15; Page 44-45; 119pp; English.  
xx  
cc The invention relates to methods for the heterologous expression of  
cc Neisserial proteins from Neisseria meningitidis and Neisseria  
cc gonorrhoeae. At least one domain in the protein is deleted, e.g. the  
cc leader peptide, and may be replaced by a domain from a different protein  
cc to make a fusion protein, in order to enhance heterologous expression of  
cc Neisserial proteins. Also, a region of a protein, such as a poly-glycine  
cc stretch, can be mutated to enhance expression. The proteins used in the  
cc processes include ORF46.1, 287, 741, 919, 953, 961 and 983. Sequences  
cc AAS43868-AAS43905 represent DNA molecules encoding Neisserial proteins  
cc and peptide regions of the invention.  
xx  
sq Sequence 3939 BP; 1046 A; 1100 C; 1107 G; 686 T; 0 other;  
  
Query Match 92.8%; Score 2971.8; DB 22; Length 3939;  
Best Local Similarity 97.2%; Pred. No. 0;  
Matches 3052; Conservative 0; Mismatches 47; Indels 42; Gaps 1;  
  
QY 103 ACTTCTGCGCCCGACTTCAATGCGAGCGCGCCGATGCGCAGCAACAGCAGCAACA 162  
Db 4 ACTTCTGCGCCCGACTTCAATGCGAGCGGTACCGGTATCGGCAGCAACAGCAGCAACA 63  
QY 163 ACAGGGAATCAGCAGCAGTATCTTACCGCGGTATCAAGAAGCAATGTGCAAGACAGA 222  
Db 64 ACAGGGAATCAGCAGCAGTATCTTACCGCGGTATCAAGAAGCAATGTGCAAGACAGA 123  
QY 223 AGCATGCTCTGCGCGTGGGATGACGTTGCGGTTACAGACAGGATGCCAAATCAAT 282  
Db 124 AGCATGCTCTGCGCGTGGGATGACGTTGCGGTTACAGACAGGATGCCAAATCAAT 183  
QY 283 GCCCCCCCCCCGAACTGCTGATACCGGAGACTTCAACACCCAAATGAGCATACAGAA 342  
Db 184 GCCCCCCCCCCGAACTGCTGATACCGGAGACTTCCAAACCCAAATGAGCATACAGAA 243  
QY 343 TTGATCAACCTCAAACTGCAATTTGAACGAGCTATACAGGACGCGGTAGAGTAGGT 402  
Db 244 TTGATCAACCTCAAACTGCAATTTGAACGAGCTATACAGGACGCGGTAGAGTAGGT 303  
QY 403 ATCGTCGATACAGGCGAATTCGTCGCGACATATCCTTTCCCGAACTGTATGGCAGAAA 462  
Db 304 ATCGTCGACACAGGCGAATTCGTCGCGACATATCCTTTCCCGAACTGTATGGCAGAAA 363  
QY 463 GAACAGCGCTATACGCAAAATACAAAACACTATACGGCGTATATCGGAAGAAAGCGCT 522  
Db 364 GAACAGCGCTATACGCAAAATACAAAACACTATACGGCGTATATCGGAAGAAAGCGCT 423
```



Qy	523	GAAGACGAGCGGTAAAGACATTAAGACCTTCTTTTCGACGATAGGCCGTTTATAGACT	582
Db	424	GAAGCGGAGCGGTAAAGACATTTGAAGCTTCTTTTCGACGATAGGCCGTTTATAGACT	483
	583	GAAGCAAAGCCGACGGATATCCGCCACGTAAAGAAATCGGACACATCGATGTGGTCTCC	642
Db	484	GAAGCAAAGCCGACGGATATCCGCCACGTAAAGAAATCGGACACATCGATTTGGTCTCC	543
Qy	643	CATATTATTCGGGGGGTTCGTGGACGCGACACTCGAGCGGTATTTCGCCCGCATGCG	702
Db	544	CATATTATTCGGGGGGTTCGTGGACGCGACACTCGAGCGGTATTTCGCCCGCATGCG	603
Qy	703	ACGCTACACATATGAATACGGATGATGCAACCAAGACGAATAATGCTCTGACGCCATC	762
Db	604	ACGCTACACATATGAATACGGATGATGAAACCAAGACGAATAATGATGGTTTCGACGCATC	663
Qy	763	CGCAATCGATGGGTCAAGCTGGCGCAACGTGGCGTGCCTACGTCGTCATTAACAGTTTTCGA	822
Db	664	CGCAATCGATGGGTCAAGCTGGCGCAACGTGGCGTGCCTACGTCGTCATTAACAGTTTTCGA	723
Qy	823	ACAACATCGAGGCGAGCGACTCGCGACCATTTTCCTAAATAGCCAAATCGGAGGACGATAC	882
Db	724	ACAACATCGAGGCGAGCGACTCGCGACCTTTTTCCTAAATAGCCAAATCGGAGGACGATAC	783
Qy	883	CGCAACGCTTGCTCCGCTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCTCATG	942
Db	784	CGCAACGCTTGCTCCGCTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCTCATG	843
Qy	943	CAACAGAGCGATTACGGCAACTTGCTACCAACATCCGTAAATAAAACATCGCTTTTCATT	1002
Db	844	CAACAGAGCGATTACGGCAACCTGCTCTACCACATCCGTATAAACATGCTTTTTCATC	903
Qy	1003	TTTTCGCAAGCAATGACGCAAGCTCAGCCCAACACATGACCCCTATTGCCCCATTTAT	1062
Db	904	TTTTCGCAAGCAATGACGCAAGCTCAGCCCAACACATATGCCCTATTGCCATTTTAT	963
Qy	1063	GAATAAGATGCTCAAAAGGCAATTATCACAGTCGCGAGCGGTAGACCCAGTCGAGAAAG	1122
Db	964	GAATAAGATGCTCAAAAGGCAATTATCACAGTCGCGAGCGGTAGACCCAGTCGAGAAAG	1023
Qy	1123	TTCA-----ATGGCTCCAACCAT	1140
Db	1024	TTCAACGGGAAATGATGGAGAACCGGTACAGAACCGCTTGAGTATGGCTCCAACCAT	1083
Qy	1141	TGCGGAAATTACTGCCATGTGGTGCTATTCGGACCCCTATGAAGCAAGGTCCTGTTCAAC	1200
Db	1084	TGCGGAAATTACTGCCATGTGGTGCTGTCGGCACCCCTATGAAGCAAGGTCCTGTTCAAC	1143
Qy	1201	CGTACAAACCCGATTCAAATTCCGGAACATCCTTTTCCGCACCCATCGCTAAACGGGACG	1260
Db	1144	CGTACAAACCCGATTCAAATTCCGGAACATCCTTTTCCGCACCCCTCGTAAACGGGACG	1203
Qy	1261	CGCGCTCTGCTGTCAGAAATACCCGTTGGATGAGCAACGACAACTGCGTACCACGCTG	1320
Db	1204	CGCGCTCTGCTGTCAGAAATACCCGTTGGATGAGCAACGACAACTGCGTACCACGCTG	1263
Qy	1321	CTGACAAACGCTCAGACATCGGTGCGCTGCGCTGAGACCAAGTTCGGCTGGGGAAGT	1380
Db	1264	CTGACAAACGCTCAGACATCGGTGCGCTGCGCTGAGACCAAGTTCGGCTGGGGAAGT	1323
Qy	1381	CTGGATGCGGGTAAGGCCATGAACGGACCGCGTCTCTTTCCGTTTCGGGACTTTTACCGCC	1440
Db	1324	CTGGATGCGGGTAAGGCCATGAACGGACCGCGTCTCTTTCCGTTTCGGGACTTTTACCGCC	1383
Qy	1441	GATACAAAGGTTACATCCGATATTGCTTACTCTTCCGTAACGACATTTTCAGGCACGGC	1500
Db	1384	GATACAAAGGTTACATCCGATATTGCTTACTCTTCCGTAACGACATTTTCAGGCACGGC	1443
Qy	1501	GGCTTGATCAAAAAGGGCGGACCACTGCACTGACGCGCAACACACCTATACGGC	1560
Db	1444	GGCTTGATCAAAAAGGGCGGACCACTGCACTGACGCGCAACACACCTATACGGC	1503



Db	664	CCCAATGCATTGGGTCAAGCTGGCGCAACCTGGCGGTGCGGATCGTCAATTAACAGTTTGGG	723
Qy	823	ACAACTCGAGGGCAGCGCACTGCCGACCAATTTCCAAATAGCAATTCGGAGGACGACTAC	882
Db	724	ACAACTCGAGGGCAGCGCACTGCCGACCTTTTCCAAATAGCCAATTCGGAGGACGACTAC	783
Qy	883	CGCCAAGCGTTGCTCGCCTATTTCGGCGGTGATAAAAACAGACGAGGGTATCCGGCTGATG	942
Db	784	CGCCAAGCGTTGCTCGCACTATTTCGGCGGTGATAAAACAGACGAGGGTATCCCGCTGATG	843
Qy	943	CAACAGACGATTACGCAACTTGTCCTACCAACATCCGTAATAAAACATGCTTTTTCATT	1002
Db	844	CAACAGACGATTACGGCAACTGTCTTACCAACATCCCGTAATAAAACATGCTTTTTCATC	903
Qy	1003	TTTTTCGCAAGCAATCAGCCACAAGCTCAGCCCAACACACTGACCCATTATGCCATTTTAT	1062
Db	904	TTTTTCGACAGCAATCAGCCACAAGCTCAGCCCAACACATATGCCCTATTGCCCATTTTAT	963
Qy	1063	GAATAAGATGCTCAAAAGGCATTATCAGTTCGACGCGTAGCCGCACTGGAGAAAAG	1122
Db	964	GAATAAGACGCTCAAAAGGCATTATCAGTTCGACGCGTAGCCGCACTGGAGAAAAG	1023
Qy	1123	TTCA-----ATGGCTCCAACCAT	1140
Db	1024	TTCAACGGGAAATGATGGAGAACGGGTACAGAACCGCTTGAGTATGGGTCCAACCAT	1083
Qy	1141	TGGGAAATTACTGCCATGTGGTGCTTATCGGCACCCCTATGAAGCAAGCTCGGTTTCACC	1200
Db	1084	TGGGAAATTACTGCCATGTGGTGCTTATCGGCACCCCTATGAAGCAAGCTCGGTTTCACC	1143
Qy	1201	CGTACAAACCCGATTCAAAATTCGGGAAACATCTTTTCCGCACCCATCGTAAACCGGCAG	1260
Db	1144	CGTACAAACCCGATTCAAAATTCGGGAAACATCTTTTCCGCACCCATCGTAAACCGGCAG	1203
Qy	1261	CGCGCTCTGCTGTCAGAAATACCGTGGATGACGACGACAACTGCGTACCACGCTG	1320
Db	1204	CGCGCTCTGCTGTCAGAAATACCGTGGATGACGACGACAACTGCGTACCACGCTG	1263
Qy	1321	CTGACAAACGGCTCAGACACATCGGTGAGTCGGCTGGACACAAGTTTCGGCTGGGACTG	1380
Db	1264	CTGACGACGGCTCAGACACATCGGTGAGTCGGCTGGACACAAGTTTCGGCTGGGACTG	1323
Qy	1381	CTGGATCGCGGTAAAGCCATGAACGGACCGCGTCCCTTTCGTCGCGACTTTACCGCC	1440
Db	1324	CTGGATCGCGGTAAAGCCATGAACGGACCGCGTCCCTTTCGTCGCGACTTTACCGCC	1383
Qy	1441	GATTACGAAAGGTACATCCGATATTGCCCTACTCCTTCCGTACGACATTTTTCAGGCACGGG	1500
Db	1384	GATTACGAAAGGTACATCCGATATTGCCCTACTCCTTCCGTACGACATTTTTCAGGCACGGG	1443
Qy	1501	GGCTGTATCAAAAAGCGCGCACCACTGCAACTGACGCAACAACACTATACGGG	1560
Db	1444	GGCTGTATCAAAAAGCGCGCACCACTGCAACTGACGCAACAACACTATACGGG	1503
Qy	1561	AAAAACATTATCGAAGCGGTTTCGCTGTGTTGTTACGGCAACAACAACTCGATATGCGC	1620
Db	1504	AAAAACATTATCGAAGCGGTTTCGCTGTGTTGTTACGGCAACAACAACTCGATATGCGC	1563
Qy	1621	GTGAAACCAAGGTCGCTGATTTATTAACGGGGCGGATCCGGCGGTACGCTGAACAGC	1680
Db	1564	GTGAAACCAAGGTCGCTGATTTATTAACGGGGCGGATCCGGCGGCACTCCGGCGCAACAGC	1623
Qy	1681	GACGGCATTGTCTATCTGGCAGATACCGACCGATCCGGCGCAACACCGGTGACACATC	1740
Db	1624	GACGGCATTGTCTATCTGGCAGATACCGACCGATCCGGCGCAACACCGGTGACACATC	1683
Qy	1741	AAAGGCGCATCTGCAGCTGGGCGGGAAGGTACGCTGTACACACACTTTTGGCAAACTGCTG	1800
Db	1684	AAAGGCGCATCTGCAGCTGGACGGCAAGGTAAGCTGTACACACTTTTGGCAAACTGCTG	1743
Qy	1801	AAAGTGCAGGTCACCGGATGACCGCGGCAAGCTGTACATGTCCGCAACCGGCAAGGG	1860
Db	1744	AAAGTGCAGGTCACCGGATGACCGCGGCAAGGTAAGCTGTACACACTTTTGGCAAACTGCTG	1803

Qy	1861	CGAGGCTATCTCAACCGGTACCGGACAAACGTGTTCCTTCCCTGAGTCGCGCCAAAATCGGG	1921
Db	1804	CGAGGCTATCTCAACAGTACCGGACGACGTGTTCCTTCTCTGAGTCGCGCCAAAATCGGG	1863
Qy	1921	CGGATTTATTTCTTTCTCAACAACATCGAAACCGAGCGGTGTCGTCTGGCTTCCCTCGAC	1980
Db	1864	CAGATTTATTTCTTTCTCAACAACATCGAAACCGAGCGCGGCTGCTGGCTTCCCTCGAC	1923
Qy	1981	AGCGTCGAAAAAACAGCGGCGAGTGAAGCGCACAGCGTGTCTCTATTATGTCCGTGCGGGC	2040
Db	1924	AGCGTCGAAAAAACAGCGGCGAGTGAAGCGCACAGCTGTCTCTATTATGTCCGTGCGGGC	1983
Qy	2041	AATCGGCACGGACTGTTCCGGCAGCGGCACATTTCCGGCCCGCGCGGTCTTGAACACGCGC	2100
Db	1984	AATCGGCACGGACTGTTCCGGCAGCGGCACATTTCCGGCCCGCGCGGTCTTGAACACGCGC	2043
Qy	2101	GTAGAACAGGGCGGCGAGCAATCTTGAAAAACCTGTATGTTGTCGAATCTGGATCGCTCGGAATCA	2160
Db	2044	GTAGAACAGGGCGGCGAGCAATCTTGAAAAACCTGTATGTTGTCGAATCTGGATCGCTCGGAATCA	2103
Qy	2161	TCCGCAACACCCGAGCGGTTGAAACTGCGCGCAGCGCGCGCACAGATATCGCGGGCATC	2220
Db	2104	TCCGCAACACCCGAGCGGTTGAAACTGCGCGCAGCGCGCACAGATATCGCGGGCATC	2163
Qy	2221	CGCCCTTACGGCGCAACTTTTCCGCGCAGCGCGCAGCGCTACAGCATGCGAATCGCGCCGAC	2280
Db	2164	CGCCCTTACGGCGCAACTTTTCCGCGCAGCGCGCAGCGCTACAGCATGCGAATCGCGCCGAC	2223
Qy	2281	GGTTAGCGATCTTCAACAGTCTCGCCGCTACCGTCTATGCGGACAGTACCGCGCCCAT	2340
Db	2224	GGTTAGCGATCTTCAACAGTCTCGCCGCTACCGTCTATGCGGACAGTACCGCGCCCAT	2283
Qy	2341	GCCGATATGACGAGCGCGCGCTGAAAGCCGTATCGGACGGGTGGACCAACAAGCTTACG	2400
Db	2284	GCCGATATGACGAGCGCGCGCTGAAAGCCGTATCGGACGGGTGGACCAACAAGCTTACG	2343
Qy	2401	GGTCTGCGCTCATCGCGCAACCCCAACAGGACGGTGGAAAGTGGAAACAGCGCGTGT	2460
Db	2344	GGTCTGCGCTCATCGCGCAACCCCAACAGGACGGTGGAAAGTGGAAACAGCGCGTGT	2403
Qy	2461	GAAGGCAAAATGCGCGCGAGTACCCAAACCGTTCGGCATTTGCCGCGAAAAACCGCGGAAAT	2520
Db	2404	GAAGGCAAAATGCGCGCGCGAGTACCCAAACCGTTCGGCATTTGCCGCGAAAAACCGCGGAAAT	2463
Qy	2521	ACGACAGCGCGCGCACATCGGCATGGGACACACACATGGAGCGGAACAGTGCAAAT	2580
Db	2464	ACGACAGCGCGCGCACATCGGCATGGGACACACACATGGAGCGGAACAGTGCAAAT	2523
Qy	2581	GCAAAAACCGACAGCATTAGTCTGTTTGAGGCATACGGCACGATGCGGGCGATATCGGC	2640
Db	2524	GCAAAAACCGACAGCATTAGTCTGTTTGAGGCATACGGCACGATGCGGGCGATATCGGC	2583
Qy	2641	TATCTAAAGCGCTGTTCTCTACGGACGCTTCAAAAAACAGCATACGCCGACGCGGT	2700
Db	2584	TATCTAAAGCGCTGTTCTCTACGGACGCTTCAAAAAACAGCATACGCCGACGCGGT	2643
Qy	2701	GCGGACCAACATCGGAAGCGAGCTACAGCGACGCTGATGCAGCTGGGCGCATCGGC	2760
Db	2644	GCGGACCAACATCGGAAGCGAGCTACAGCGACGCTGATGCAGCTGGGCGCATCGGC	2703
Qy	2761	GGTGCAACGTTCCGTTTTCGCGCAACGGGAGATTTGACGGTTCGAAGCGGTCTCGCGTAC	2820
Db	2704	GGTGCAACGTTCCGTTTTCGCGCAACGGGAGATTTGACGGTTCGAAGCGGTCTCGCGTAC	2763
Qy	2821	GACCTGCTCAAACAGGATGCATTCGCCGAAAAAGCAGTGTCTTTGGCTGGAGCGGCAAC	2880
Db	2764	GACCTGCTCAAACAGGATGCATTCGCCGAAAAAGCAGTGTCTTTGGCTGGAGCGGCAAC	2823
Qy	2881	AGCCTACTGAAGGCGACATGTCGCGAGCTCGCGGGTCTGAAGCTGTTCGCAACCTTTGAGC	2940
Db	2824	AGCCTACTGAAGGCGACGTCGTCGAGCTTCGCGGGTCTGAAGCTGTTCGCAACCTTTGAGC	2883

QY 2941 GATAAAGCGCTCCTGTTTCAACAGCGCGCGGTGGAAACGCGACCTGAACGAGCGCGACTAC 3000  
Db 2884 GATAAAGCGCTCCTGTTTCAACAGCGCGCGGTGGAAACGCGACCTGAACGAGCGCGACTAC 2943  
QY 3001 ACGGTAACGGCGCGCTTTACCGCGCGGACTGAGCAACCGGCAAGACGGGGCGACGCAAT 3060  
Db 2944 ACGGTAACGGCGCGCTTTACCGCGCGGACTGAGCAACCGGCAAGACGGGGCGACGCAAT 3003  
QY 3061 ATGCCGCGACACCGCGCTGTTCCGCTGTGGCGGCGATGTCGAATTCGGCAACGGCTGG 3120  
Db 3004 ATGCCGCGACACCGCGCTGTTCCGCTGTGGCGGCGATGTCGAATTCGGCAACGGCTGG 3063  
QY 3121 AACGGCTTGGCAGCTTACAGCTACGCGCGGTTCCAAACAGTACGGCAACACAGCGGACGA 3180  
Db 3064 AACGGCTTGGCAGCTTACAGCTACGCGCGGTTCCAAACAGTACGGCAACACAGCGGACGA 3123  
QY 3181 CTCGGGCTAGGCTACGGGTTTC 3201  
Db 3124 GTCGGGCTAGGCTACGGGTTTC 3144

RESULT 14  
AAS43877  
ID AAS43877 standard; DNA; 4179 BP.  
XX AAS43877;  
AC  
XX  
XX  
DT 18-DEC-2001 (first entry)  
XX  
DE Neisseria meningitidis fusion protein delta-G983-961c DNA.  
KW Neisseria gonorrhoeae; leader peptide; fusion protein; ORF46.1; ds;  
KW Neisserial protein.  
OS Neisseria meningitidis.  
OS Synthetic.  
XX  
XX WO200164922-A2.  
XX  
XX 07-SEP-2001.  
XX  
XX 28-FEB-2001; 2001WO-IB00452.  
XX  
XX 28-FEB-2000; 2000GB-0004695.  
XX 13-NOV-2000; 2000GB-0027675.  
XX  
XX (CHIR-) CHIRON SPA.  
XX  
XX Arico MB, Comanducci M, Galeotti C, Massignani V, Guilianani MM;  
PI Pizza M;  
XX  
XX WPI; 2001-582163/65.  
DR P-PSDB; AAU27578.  
XX  
XX Producing heterologous proteins from Neisseria meningitidis and N.  
PT gonorrhoeae -  
XX  
XX Example 15; Page 47-49; 119pp; English.  
XX  
XX The invention relates to methods for the heterologous expression of  
CC Neisserial proteins from Neisseria meningitidis and Neisseria  
CC gonorrhoeae. At least one domain in the protein is deleted, e.g. the  
CC leader peptide, and may be replaced by a domain from a different  
CC to make a fusion protein, in order to enhance heterologous expression of  
CC Neisserial proteins. Also, a region of a protein, such as a poly-glycine  
CC stretch, can be mutated to enhance expression. The proteins used in the  
CC processes include ORF46.1, 287, 741, 919, 953, 961 and 983. Sequences  
CC AAS43868-AAS43905 represent DNA molecules encoding Neisserial proteins  
CC and peptide regions of proteins of the invention.  
XX  
SQ Sequence 4179 BP; 1187 A; 1140 C; 1121 G; 731 T; 0 other;  
Query Match 92.8%; Score 2971.8; DB 22; Length 4179;

Best Local Similarity 97.2%; Pred. No. 0;  
Matches 3052; Conservative 0; Mismatches 47; Indels 42; Gaps 1;  
QY 103 ACTTCTGCGCCCGACTTCAATGCGAGCGCGCCGATATCGGAGCAACAGCAGACGACACA 162  
Db 4 ACTTCTGCGCCCGACTTCAATGCGAGCGCGTACCGGTATCGGAGCAACAGCAGACGACACA 63  
QY 163 ACAGCGAAATACGAGCAGTATCTTACGCGGTATCAAGAAGCAAAATGTGCAAGACAGA 222  
Db 64 ACAGCGAAATACGAGCAGTATCTTACGCGGTATCAAGAAGCAAAATGTGCAAGACAGA 123  
QY 223 AGCATGCTCTGTGCGGTGCGGATGACGTGCGGTATCAGACAGGATGCCAAAATCAAT 282  
Db 124 AGCATGCTCTGTGCGGTGCGGATGACGTGCGGTATCAGACAGGATGCCAAAATCAAT 183  
QY 283 GCCCGCCCGCCCAATCTGCATACCGGAGACTTTACAAACCCCAATGACGCATACAGAAT 342  
Db 184 GCCCGCCCGCCCAATCTGCATACCGGAGACTTTCAAACCCCAATGACGCATACAGAAT 243  
QY 343 TTGATCAACCTCAAACTGCAATTGAAGCAGGCTATACAGGACGCGGGTAGAGTAGT 402  
Db 244 TTGATCAACCTCAAACTGCAATTGAAGCAGGCTATACAGGACGCGGGTAGAGTAGT 303  
QY 403 ATCTGCTGATACAGCGCAATCCGTGCGGACGATATCCTTCCCGAACTGTATGGCAGAAAA 462  
Db 304 ATCTGCTGATACAGCGCAATCCGTGCGGACGATATCCTTCCCGAACTGTATGGCAGAAAA 363  
QY 463 GNAACGCGCTAATACGAAAAATACAAAACTATACGGGTATATCGGAGAAAGCGCT 522  
Db 364 GNAACGCGCTAATACGAAAAATACAAAACTATACGGGTATATCGGAGAAAGCGCT 423  
QY 523 GAAGACGAGGCGGTAAAGACATTAAGCTTCTTCGACGATGAGCGCTTATAGAGACT 582  
Db 424 GAAGACGAGGCGGTAAAGACATTAAGCTTCTTCGACGATGAGCGCTTATAGAGACT 483  
QY 583 GAAGCAAAAGCGGATATCCGCCACGTAAAGAAATCGGACACATCGATGTGCTCTCC 642  
Db 484 GAAGCAAAAGCGGATATCCGCCACGTAAAGAAATCGGACACATCGATGTGCTCTCC 543  
QY 643 CATATTATTGGCGGCGGTCCGTGGACGCGACACCTGAGGCGGTATTCGCGCCCATCGG 702  
Db 544 CATATTATTGGCGGCGGTCCGTGGACGCGACACCTGAGGCGGTATTCGCGCCCATCGG 603  
QY 703 ACGCTACACATAATGAATACGATGATGAAACCAAGAAATATGTCTCGACGCCATC 762  
Db 604 ACGCTACACATAATGAATACGATGATGAAACCAAGAAATATGTCTCGACGCCATC 663  
QY 763 CGCAATGATGGGTCAAGCTGGCGGAACTGCGGTGCGCATCGTCAATAACAGTTTTTGA 822  
Db 664 CGCAATGATGGGTCAAGCTGGCGGAACTGCGGTGCGCATCGTCAATAACAGTTTTTGA 723  
QY 823 ACAACATCGAGGCGGACCTGCGGACCATTTCCAAATAGCCAAATTCGGAGGAGCAGTAC 882  
Db 724 ACAACATCGAGGCGGACCTGCGGACCATTTCCAAATAGCCAAATTCGGAGGAGCAGTAC 783  
QY 883 CGCCAAGCGTTGCTCGCTATTCGCGCGGTGATAAAACAGAGCGGTATCGCGCTGATG 942  
Db 784 CGCCAAGCGTTGCTCGCTATTCGCGCGGTGATAAAACAGAGGATATCGCGCTGATG 843  
QY 943 CAACAGCGATTAACGCAACTGTCTTACACATATCCGTATTAATAACATGCTTTTCATT 1002  
Db 844 CAACAGCGATTAACGCAACTGTCTTACACATATCCGTATTAATAACATGCTTTTCATT 903  
QY 1003 TTTTCGGCAAGCAATGAGCACAAGCTGAGCCCAACACTGACCCCTATTCGCAATTTAT 1062  
Db 904 TTTTCGGCAAGCAATGAGCACAAGCTGAGCCCAACACTGACCCCTATTCGCAATTTAT 963  
QY 1063 GAAAAAGATGCTCAAAAAGGCAATATCAGTCGCGGCTAGACCGAGTCGAGGAGAAAG 1122  
Db 964 GAAAAAGATGCTCAAAAAGGCAATATCAGTCGCGGCTAGACCGAGTCGAGGAGAAAG 1023  
QY 1123 TTCA-----ATGGCTCCAAACCAT 1140  
|||||

Db 1024 TTCAAACGGGAAATGTATGGAGAACCGGGTACAGAACCGCTTGAGTATGGCTCCAAACCAT 1083  
QY 1141 TCGGGAATTACTGCCATGTGGTCCCTATCGGCACCCCTATGAGCAAGCGGTCCGTTTTCACC 1200  
Db 1084 TCGGGAATTACTGCCATGTGGTCCCTGTGCGCACCCCTATGAGCAAGCGGTCCGTTTTCACC 1143  
QY 1201 CGTACAAACCCGATTCAAATTTGCGGAAGCATCTTTTCGCGACCCCATCGTTAAACCGGCAGG 1260  
Db 1144 CGTACAAACCCGATTCAAATTTGCGGAACATCCTTTTCGCGACCCCATCGTTAAACCGGCAGG 1203  
QY 1261 GGGGCTCTGCTGTCAGAGAAATACCCGTGTAGAGAGCAAGCAACACCTGCGTACCAACGCTG 1320  
Db 1204 GCGGCTCTGCTGTCAGAGAAATACCCGTGTAGAGAGCAAGCAACACCTGCGTACCAACGCTG 1263  
QY 1321 CTGACAAACGGCTCAGACACATCGGTGAGTCGGGTGTGAGAGCAAGTTTCGGCTGGGACTG 1380  
Db 1264 CTGACAAACGGCTCAGACACATCGGTGAGTCGGGTGTGAGAGCAAGTTTCGGCTGGGACTG 1323  
QY 1381 CTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTTCGGCGACTTTTACCGCC 1440  
Db 1324 CTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTTCGGCGACTTTTACCGCC 1383  
QY 1441 GATACAAAGGTACATCCGATATGGCTACTCTTCGTTAAACACATTTTACGGACGCGG 1500  
Db 1384 GATACAAAGGTACATCCGATATGGCTACTCTTCGTTAAACACATTTTACGGACGCGG 1443  
QY 1501 GGCTGATCAAAAAGCGCGCAGCAACTGCAACTGCAGGCAACCAACACCTATACGGG 1560  
Db 1444 GGCTGATCAAAAAGCGCGCAGCAACTGCAACTGCAGGCAACCAACACCTATACGGG 1503  
QY 1561 AAAACCATTTATCGAAGCGCGTTCCGTTGGTGTGTACGGCAACCAACCAAAATCGGATATCGCG 1620  
Db 1504 AAAACCATTTATCGAAGCGCGTTCCGTTGGTGTGTACGGCAACCAACCAAAATCGGATATCGCG 1563  
QY 1621 GTCGAAACCAAGGTGCGTGTATTTATACGGGGCGGCGATPCGGCGGTAGCCTGGAACAGC 1680  
Db 1564 GTCGAAACCAAGGTGCGTGTATTTATACGGGGCGGCGATPCGGCGGTAGCCTGGAACAGC 1623  
QY 1681 GACGGCATTTGCTATCTGCGAGATACCGGACGATCCGGCGCAACCAACCGGTGCACATC 1740  
Db 1624 GACGGCATTTGCTATCTGCGAGATACCGCAATCCGGCGCAACCAACCGGTGCACATC 1683  
QY 1741 AAAGCGCATCTGAGCTGGCGGCGAAGGTACGCTGTACACAGCTTTTGGGCAAACTGCTG 1800  
Db 1684 AAAGCGCATCTGAGCTGGAGCGCAAGGTACGCTGTACACAGCTTTTGGGCAAACTGCTG 1743  
QY 1801 AAAGTGAGCGGTACGGCGATGACCGGCGCAAGCTGTACATGTCCGACGCGGCAAGGG 1860  
Db 1744 AAAGTGAGCGGTACGGCGATTTATCCGGCGCAAGCTGTACATGTCCGACGCGGCAAGGG 1803  
QY 1861 GCAGGCTATCTCAACCGTACCGGACACGTTTCCCTTCTCTGAGTCCCGCAAAATCGGG 1920  
Db 1804 GCAGGCTATCTCAACAGTACCGGACGAGTGTTCCTTCTCTGAGTCCCGCAAAATCGGG 1863  
QY 1921 CGGGATTATTTCTTCTCAAAACATCGAAACCGAGGTTGCTGCTGGCTTCCCTCGAC 1980  
Db 1864 CAGGATTATTTCTTCTCAAAACATCGAAACCGAGGCGGCTGCTGGCTTCCCTCGAC 1923  
QY 1981 AGCGTGA AAAACAGCGGCGAGTGAAGGGACACGCTGTCTTATTTATGTCCTGCGGGC 2040  
Db 1924 AGCGTGA AAAACAGCGGCGAGTGAAGGGACACGCTGTCTTATTTATGTCCTGCGGGC 1983  
QY 2041 AATGGCGCAGGACTGCTTCGGCAGGGCACATTCGCGCCCGCGCTCGAACAACGCG 2100  
Db 1984 AATGGCGCAGGACTGCTTCGGCAGGGCACATTCGCGCCCGCGCTCGAACAACGCG 2043  
QY 2101 GTAGAACAGGGCGGCGAGCAATCTGGA AAACCTGATGGTTCGAACCTGCTCCGAATCA 2160  
Db 2044 GTAGAACAGGGCGGCGAGCAATCTGGA AAACCTGATGGTTCGAACCTGCTCCGAATCA 2103  
QY 2161 TCCGCAACCCCGAGACGGTTGAAACTGCGGCGCGGACCGCACAGATATGCGCGGCATC 2220  
Db 2104 TCCGCAACCCCGAGACGGTTGAAACTGCGGCGCGGACCGCACAGATATGCGCGGCATC 2163

QY 2221 CGCCCTTACGGCGCAACTTTTCCGCGCAGCGGACCGGTACAGCATGCGAATGCGCGCAC 2280  
Db 2164 CGCCCTTACGGCGCAACTTTTCCGCGCAGCGGACCGGTACAGCATGCGAATGCGCGCAC 2223  
QY 2281 GGTGTACGCATCTTCAACAGTCTCGCGGTACCGTCTATGCCGACAGTACCGCGCGCCAT 2340  
Db 2224 GGTGTACGCATCTTCAACAGTCTCGCGGTACCGTCTATGCCGACAGTACCGCGCGCCAT 2283  
QY 2341 GCCGATATGCAGGAGCGCGGCTGAAAGCGGTATCGGACGGGTTGGACCAACAGCTACG 2400  
Db 2284 GCCGATATGCAGGAGCGCGGCTGAAAGCGGTATCGGACGGGTTGGACCAACAGCGCAC 2343  
QY 2401 GGTCTCGCGTCTATCGCGAAACCCAAACAGAGCGGTGGAACGTGGGAACAGGCGGTGT 2460  
Db 2344 GGTCTCGCGTCTATCGCGAAACCCAAACAGAGCGGTGGAACGTGGGAACAGGCGGTGT 2403  
QY 2461 GAAGGCAAAATGCGCGGCGAGTACCCAAACCGTCGGCATGCCCGGAAAAACCGCGGCAAA 2520  
Db 2404 GAAGGCAAAATGCGCGGCGAGTACCCAAACCGTCGGCATGCCCGGAAAAACCGCGGCAAA 2463  
QY 2521 ACCACAGCAGCCGCCACACTGGGCATGGGACACAGACACATGGAGCGGAAAAACAGTGCA 2580  
Db 2464 ACACAGCAGCCGCCACACTGGGCATGGGACAGCGACACATGGAGCGGAAAAACAGTGCA 2523  
QY 2581 GCAAAAACCGACAGCATTTAGTCTGTTTGCAGGCATACGGCACGATGGGCGGATATCGGC 2640  
Db 2524 GCAAAAACCGACAGCATTTAGTCTGTTTGCAGGCATACGGCACGATGGGCGGATATCGGC 2583  
QY 2641 TATCTCAAAGGCGTGTCTCTACGGACGCTACAAAACAGCATACGCCGACGACCGCGT 2700  
Db 2584 TATCTCAAAGGCGTGTCTCTACGGACGCTACAAAACAGCATACGCCGACGACCGCGT 2643  
QY 2701 GCGGAGAACATCGGAGGACGCTCAACGGCACGCTGATGACGCTGGGCGGACCTGGGC 2760  
Db 2644 GCGGAGAACATCGGAGGACGCTCAACGGCACGCTGATGACGCTGGGCGGACCTGGGC 2703  
QY 2761 GGTGTCAACGTTCCGTTTGGCGCAACGGGAGATTTGACGCTCGAAGCGCTCTGCGGTAC 2820  
Db 2704 GGTGTCAACGTTCCGTTTGGCGCAACGGGAGATTTGACGCTCGAAGCGCTCTGCGGTAC 2763  
QY 2821 GACCTGCTCAACAGGATGATTCGCCGAAAAAGGACGCTTTTGGGCTGGAGCGGCAAC 2880  
Db 2764 GACCTGCTCAACAGGATGATTCGCCGAAAAAGGACGCTTTTGGGCTGGAGCGGCAAC 2823  
QY 2881 AGCCTACTGAAGCACACTGGTTCGAGCTCGGGGCTGAAGCTGTGCGCAACCCCTTGAGC 2940  
Db 2824 AGCCTACTGAAGCACACTGGTTCGAGCTCGGGGCTGAAGCTGTGCGCAACCCCTTGAGC 2883  
QY 2941 GATAAAGCCGCTCCTGTTTGAACGGGCGGTGGAACGCGACCTGAAACGACGCGGACTAC 3000  
Db 2884 GATAAAGCCGCTCCTGTTTGAACGGGCGGTGGAACGCGACCTGAAACGACGCGGACTAC 2943  
QY 3001 ACGGTAACGGGCGCTTTTACCGCGGACTGCAACACCGGCAAGAGCGGGGCGACGCAAT 3060  
Db 2944 ACGGTAACGGGCGCTTTTACCGCGGACTGCAACACCGGCAAGAGCGGGGCGACGCAAT 3003  
QY 3061 ATGCGCACACCCGCTGTTTGGGCTCGGGGCGGATGTCGAAATTCGCGCAACGGGCTG 3120  
Db 3004 ATGCGCACACCCGCTGTTTGGGCTCGGGGCGGATGTCGAAATTCGCGCAACGGGCTG 3063  
QY 3121 AAGCGCTTGGCAGCTTACAGCTACGCGGTTTCCAAACAGTACGCAACACAGCGGACGA 3180  
Db 3064 AAGCGCTTGGCAGCTTACAGCTACGCGGTTTCCAAACAGTACGCAACACAGCGGACGA 3123  
QY 3181 GTCGGCGTAGGCTACCGGTTTC 3201  
Db 3124 GTCGGCGTAGGCTACCGGTTTC 3144

RESULT 15  
AADI7042  
ID AADI7042 standard; DNA; 4179 BP.

```
XX AAD17042;
AC
XX
XX 29-NOV-2001 (first entry)
DT
XX
XX N. meningitidis strain 2996 delta G983-961c fusion DNA.
DE
XX
XX Heterologous expression; Neisserial protein;
KW delta G983-961c fusion protein; ds.
KW
XX
XX Neisseria meningitidis 2996.
OS
XX
XX Key Location/Qualifiers
FH 1..4179
CDS /*tag= a
FT /product= "N. meningitidis strain 2996 delta
FT G983-961c fusion protein"
FT
XX
XX WO200164920-A2.
PN
XX
XX 07-SEP-2001.
PD
XX
XX 28-FEB-2001; 2001WO-IB00420.
PF
XX
XX 28-FEB-2000; 2000GB-0004695.
PR 13-NOV-2000; 2000GB-0027675.
PR
XX
XX (CHIR-) CHIRON SPA.
PA
XX
XX Arico MB, Comanducci M, Galeotti C, Massignani V, Giuliani MM;
PI Pizza M;
PI
XX
XX WPI; 2001-557776/62.
DR
XX
XX P-PSDB; AAE10025.
DR
XX
XX Heterologous expression for the expression of two or more Neisserial
PT proteins in fused state -
PT
XX
XX Example 3; Page 17-18; 52pp; English.
PS
XX
XX The present invention relates to a method for simultaneous heterologous
CC expression of two or more Neisserial proteins which are in a fused
CC state. The method is useful for simultaneous heterologous expression of
CC two or more Neisserial proteins. A protein that may be unstable or
CC poorly expressed on its own is assisted by adding a suitable hybrid
CC partner and commercial manufacture is simplified-only one expression and
CC purification need to be employed in order to produce two separately-
CC useful proteins. The present sequence is a DNA encoding
CC Neisseria meningitidis (serogroup B, strain 2996) delta G983-961c
CC fusion protein.
XX
XX Sequence 4179 BP; 1187 A; 1140 C; 1121 G; 731 T; 0 other;
SQ

Query Match 92.8%; Score 2971.8; DB 22; Length 4179;
Best Local Similarity 97.2%; Pred. No. 0;
Matches 3052; Conservative 0; Mismatches 47; Indels 42; Gaps 1;

QY 103 ACTTCTGCGCCGACTTCAATGCAGCGCCACCGCTATCGGAGCAACAGCAGACACA 162
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 4 ACTTCTGCGCCGACTTCAATGCAGCGGTACCGGTATCGGAGCAACAGCAGACACA 63
QY 163 ACAGCGAAATCAGCAGCAGTATCTTACGCGCGGTATCAAGAAGCAATGTGCAAGACAGA 222
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 64 ACAGCGAAATCAGCAGCAGTATCTTACGCGGTATCAAGAAGCAATGTGCAAGACAGA 123
QY 223 AGCATGCTCTGTGCGGTGCGGATCAGCTTGGGTTACAGACAGGATGCCAAATCAAT 282
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 124 AGCATGCTCTGTGCGGTGCGGATCAGCTTGGGTTACAGACAGGATGCCAAATCAAT 183
QY 283 GCCCCCCCCCCCAATTCGATACCGGAGACTTTTCAAAACCCCAATGAGCATACAAGAT 342
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 184 GCCCCCCCCCCCAATTCGATACCGGAGACTTTTCAAAACCCCAATGAGCATACAAGAT 243
.
```

```
QY 343 TTGATCAACCTCAACCTGCAATTGAAGCAGGCTATACAGACGCGGGGTAGAGTAGGT 402
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 244 TTGATCAACCTCAACCTGCAATTGAAGCAGGCTATACAGACGCGGGGTAGAGTAGGT 303
QY 403 ATCGTTCGATACAGGGAATCCGTGGCAGCATATCCTTTCCGAACTGTATGGCAGAAAA 462
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 304 ATCGTTCGACACAGGGAATCCGTGGCAGCATATCCTTTCCGAACTGTATGGCAGAAAA 363
QY 463 GAACACGGCTATAACGAAAAATTACAAAAACTATACGGCGTATATCGGGAAGAACGCGCT 522
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 364 GAACACGGCTATAACGAAAAATTACAAAAACTATACGGCGTATATCGGGAAGAACGCGCT 423
QY 523 GAAGACGGAGCGGTAAAGACATTAAGCTTCTTTTCGAGATGAGCGCGTTATAGAGACT 582
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 424 GAAGACGGAGCGGTAAAGACATTAAGCTTCTTTTCGAGATGAGCGCGTTATAGAGACT 483
QY 583 GAACAAAGCCGACGGATATCCGCCACTTAAAGAAATCGGACACATCGATGTGCTTCC 642
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 484 GAACAAAGCCGACGGATATCCGCCACTTAAAGAAATCGGACACATCGATGTGCTTCC 543
QY 643 CATATTATTGGCGGCTTCCGTGGACGGCAGACTCTGCAGCGGTATTTCCGCCCGCATCG 702
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 544 CATATTATTGGCGGCTTCCGTGGACGGCAGACTCTGCAGCGGTATTTCCGCCCGCATCG 603
QY 703 ACGTACACATAATGAATACGATGATGAACCAAGAACGAAATAATGTCTGCAGCCATC 762
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 604 ACGTACACATAATGAATACGATGATGAACCAAGAACGAAATAATGTCTGCAGCCATC 663
QY 763 CGCAATGATGGTCAAGCTGGCGAAGCTGGCGTCCCATCGTCAATAAAGATTTTGA 822
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 664 CGCAATGATGGTCAAGCTGGCGAAGCTGGCGTCCCATCGTCAATAAAGATTTTGA 723
QY 823 ACAACATCGAGGCGGACCTGCCGACCATTTCCAAATAGCCAAATTCGGAGGAGCAGTAC 882
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 724 ACAACATCGAGGCGGACCTGCCGACCTTTCCAAATAGCCAAATTCGGAGGAGCAGTAC 783
QY 883 CGCAAGGCTTGTCTGCTATTTCCGGCGGTATGAACAGAGGAGGTATCGGCTGATG 942
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 784 CGCAAGGCTTGTCTGCTATTTCCGGCGGTATGAACAGAGGAGGTATCGGCTGATG 843
QY 943 CAACAGCGGATAGCGCAACTTGTCTACACATCCGCTAATAAAACATGCTTTTCATT 1002
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 844 CAACAGCGGATAGCGCAACTGTCCTACACATCCGCTAATAAAACATGCTTTTCATT 903
QY 1003 TTTTCGGAAGCAATGAGCACAAGCTCAGCCCAACACACTGACCTATTGCCATTTTAT 1062
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 904 TTTTCGAGAGCAATGAGCACAAGCTCAGCCCAACACATATGCCCTATTGCCATTTTAT 963
QY 1063 GAAAAGATGCTCAAAAAGGCAATATACAGTCGAGCGGTAGACCGCAGTGGAGAAAAG 1122
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 964 GAAAAGAGCGCTCAAAAAGGCAATATACAGTCGAGCGGTAGACCGCAGTGGAGAAAAG 1023
QY 1123 TTCA-----ATGGCTCCAACCAT 1140
DB |||||
DB 1024 TTCAACGGGAATGTATGGAGAACCGGTACAGAACCGCTTGTAGTATGGCTCCAACCAT 1083
QY 1141 TCGGAATTAATCTGCGCATGTGGTCCCTATCGGCACCTATGAAGCAAGGCTCGGTTTCACC 1200
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 1084 TCGGAATTAATCTGCGCATGTGGTCCCTATCGGCACCTATGAAGCAAGGCTCGGTTTCACC 1143
QY 1201 CGTACAAACCGGATTCAAATTTGCCGGAACATCCTTTTCCGACCCCATCGTAACCGGACG 1260
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 1144 CGTACAAACCGGATTCAAATTTGCCGGAACATCCTTTTCCGACCCCATGTAAACCGGACG 1203
QY 1261 CGGCTCTGCTCGCAGAAATACCCGTGGATGACACAGCAACCTCGGTACACGCTG 1320
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 1204 CGGCTCTGCTCGCAGAAATACCCGTGGATGACACAGCAACCTCGGTACACGCTG 1263
QY 1321 CTGACACGGCTCAGGACATCGGTGCGGTGGAGCAGCAAGTTGCGGTGGGACTG 1380
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 1264 CTGACGACGGCTCAGGACATCGGTGCGGTGGAGCAGCAAGTTGCGGTGGGACTG 1323
QY 1381 CTGGATGCGGGTAAAGCCATGAACGACCGCGGTCTTTTCCGTCGCGGACTTTTACCGCC 1440
```

|||||  
Db 1324 CTGGATCGGGTAAGGCCATGACGACCGCGTCCCTTCCGTCGGCGACTTTACGGCC 1383  
QY 1441 GATACGAAAGGTATATCCCGATATTGGCTACTCTTCGCTAAACGACATTTTCAGGCACGGGC 1500  
Db 1384 GATACGAAAGGTATATCCCGATATTGGCTACTCTTCGCTAAACGACATTTTCAGGCACGGGC 1443  
QY 1501 GGCTGATCAAAAAGCGCGCAGCCAACTGCAACTGCACGGCAACAACACTATACGGGC 1560  
Db 1444 GGCTGATCAAAAAGCGCGCAGCCAACTGCAACTGCACGGCAACAACACTATACGGGC 1503  
QY 1561 AAACCATATATCAAGCGGTTCCGCTGGTGTGTAGGCAACAACAAATCGGATATGCGC 1620  
Db 1504 AAACCATATATCAAGCGGTTCCGCTGGTGTGTAGGCAACAACAAATCGGATATGCGC 1563  
QY 1621 GTCGAAACCAAGTTCGCTGATTATTAACGGGGCGGCATCCGCGGTAGCCCTGAACAGC 1680  
Db 1564 GTCGAAACCAAGTTCGCTGATTATTAACGGGGCGGCATCCGCGGTAGCCCTGAACAGC 1623  
QY 1681 GACGGCATTTGCTATCTGGCAGATACCGACCGATCCGGCGCAACGAAACCGTACACATC 1740  
Db 1624 GACGGCATTTGCTATCTGGCAGATACCGACCAATCCGGCGCAACGAAACCGTACACATC 1683  
QY 1741 AAGGGCATCTGACGCTGGCGGGAAGGTACGCTGTACACAGTTTTGGGCAAACTGCTG 1800  
Db 1684 AAGGGCATCTGACGCTGGAGCGCAAGGTACGCTGTACACAGTTTTGGGCAAACTGCTG 1743  
QY 1801 AAAGTGCAGCGGTACGCGATACCGCGGCAAGCTGTACATGTCGGCACGCGCAAGGG 1860  
Db 1744 AAAGTGCAGCGGTACGCGATATTCGGGGCAAGCTGTACATGTCGGCACGCGCAAGGG 1803  
QY 1861 GCAGGCTATCTCAACCGTACCGGACACGCTGTTCCCTTCTCGAGTCGCGCAAAATCGGG 1920  
Db 1804 GCAGGCTATCTCAACAGTACCGGACGACGTGTTCCCTTCTCGAGTCGCGCAAAATCGGG 1863  
QY 1921 CGGGATTATTTCTTCTCAACATCTGAAACCGACGCGTGGTCTGCTGGCTTCCCTCGAC 1980  
Db 1864 CAGGATTATTTCTTCTCAACATCTGAAACCGACGCGCGCTGCTGCTTCCCTCGAC 1923  
QY 1981 AGCGTCGAAAAACAGCGGCGCTGAAGCGACACGCTGCCTATTATGTCGCTCGCGGC 2040  
Db 1924 AGCGTCGAAAAACAGCGGCGCTGAAGCGGACACGCTGCCTATTATGTCGCTCGCGGC 1983  
QY 2041 AATCGCGCAGGACTGCTTCGGCAGCGGCACATTCGCGCGCCGCGCTGTGAACACAGCC 2100  
Db 1984 AATCGCGCAGGACTGCTTCGGCAGCGGCACATTCGCGCGCCGCGCTGTGAACACAGCC 2043  
QY 2101 GTAGAACAGGCGGCGAGCAATCTGGAACCTGATGGTTCGAACTGGATGCTCCGAATCA 2160  
Db 2044 GTAGAACAGGCGGCGAGCAATCTGGAACCTGATGGTTCGAACTGGATGCTCCGAATCA 2103  
QY 2161 TCCGCAACACCCGAGCGGTTGAACTGCGCGCGCGACCGACAGATATGCGCGGCATC 2220  
Db 2104 TCCGCAACACCCGAGAGGTTGAACTGCGGCGAGCGACCGACAGATATGCGCGGCATC 2163  
QY 2221 CGCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCGGTACAGCATGCGAATGCGCGCGAC 2280  
Db 2164 CGCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCGGTACAGCATGCGAATGCGCGCGAC 2223  
QY 2281 GGTGTAGCATCTTCAACAGTCTCGCGCTACCGTCTATGCGGACAGTACCGCGGCCAT 2340  
Db 2224 GGTGTAGCATCTTCAACAGTCTCGCGCTACCGTCTATGCGGACAGTACCGCGGCCAT 2283  
QY 2341 GCGGATATGACGAGCGCGGCTGAAGCGGTATCGGACGGGTTGGACCAACAGCTACG 2400  
Db 2284 GCGGATATGACGAGCGCGGCTGAAGCGGTATCGGACGGGTTGGACCAACAGCTACG 2343  
QY 2401 GGTCTCGGGCTCATCGCGAAACCAACAGGACGCTGGAACGTTGGGAACAGGCGGTGTT 2460  
Db 2344 GGTCTCGGGCTCATCGCGAAACCAACAGGACGCTGGAACGTTGGGAACAGGCGGTGTT 2403  
QY 2461 GAAGGCAAAATCGCGGCGAGTACCCCAACCGTCGGCATTTGCCCGCAAAACCGCGCAAAAT 2520  
|||||

Db 2404 GAAGGCAAAATCGCGGCGAGTACCCAAACCGTCGGCATTTGCCGCAAAAACCGCGCAAAAT 2463  
QY 2521 ACGACAGCAGCCGCGCACCTGGCATGGGACACAGACACATGGAGCGAAAAACAGTGAAT 2580  
Db 2464 ACGACAGCAGCCGCGCACCTGGCATGGGACGAGACACATGGAGCGAAAAACAGTGAAT 2523  
QY 2581 GCAAAAACCGCACAGCATTTAGTCTGTGTTGAGGCATACGGCACGATGCGGCGATATCGGC 2640  
Db 2524 GCAAAAACCGCACAGCATTTAGTCTGTGTTGAGGCATACGGCACGATGCGGCGATATCGGC 2583  
QY 2641 TATCTCAAAAGGCTGTCTCTACGGACGCTCAAAAAACAGCATCAGCCGAGCACCGGT 2700  
Db 2584 TATCTCAAAAGGCTGTCTCTACGGACGCTCAAAAAACAGCATCAGCCGAGCACCGGT 2643  
QY 2701 GCGGACCAACATCGGGAAGGACGCTCAACGCGACGCTGATGAGCTGGGCGCACTGGC 2760  
Db 2644 GCGGACCAACATCGGGAAGGACGCTCAACGCGACGCTGATGAGCTGGGCGCACTGGC 2703  
QY 2761 GGTGTCAACGTTCCGTTTTCGCAACGGGAGATTTGACGGTTCGAAGCGGTCTGCGCTAC 2820  
Db 2704 GGTGTCAACGTTCCGTTTTCGCAACGGGAGATTTGACGGTTCGAAGCGGTCTGCGCTAC 2763  
QY 2821 GACCTGCTCAACAGAGTATTCGCCGAAAAAGGCACTGCTTTGGCTGAGGGGCAAC 2880  
Db 2764 GACCTGCTCAACAGAGTATTCGCCGAAAAAGGCACTGCTTTGGCTGAGGGGCAAC 2823  
QY 2881 AGCTCTACTCAAGGACACACTGGTCGGACTCGCGGCTCTGAAGCTGTGCGAACCTTGAGC 2940  
Db 2824 AGCTCTACTCAAGGACACCTGGTCGGACTCGCGGCTCTGAAGCTGTGCGAACCTTGAGC 2883  
QY 2941 GATAAAGCCGCTCTGTTTGAACGGCGGGGTGGAACGCGACCTGAAACGCGGCGACTAC 3000  
Db 2884 GATAAAGCCGCTCTGTTTGAACGGCGGGGTGGAACGCGACCTGAAACGCGGCGACTAC 2943  
QY 3001 ACGGTAAAGCGGCGGCTTTTACCGGCGGCGACTGAGCAACCGGCAAGCGGGGCGACGCAAT 3060  
Db 2944 ACGGTAAAGCGGCGGCTTTTACCGGCGGCGACTGAGCAACCGGCAAGCGGGGCGACGCAAT 3003  
QY 3061 ATCCGCGACACCGCGCTGTTGCGGCTCGGCGCGGATGCGAATTCGCAACGCGGTGG 3120  
Db 3004 ATCCGCGACACCGCGCTGTTGCGGCTCGGCGCGGATGCGAATTCGCAACGCGGTGG 3063  
QY 3121 AAGCGCTTGGCAGCTTACAGCTACGCGGTTCCAAACAGTACGCGCAACCGGCGGCGACGA 3180  
Db 3064 AAGCGCTTGGCAGCTTACAGCTACGCGGTTCCAAACAGTACGCGCAACCGGCGGCGACGA 3123  
QY 3181 GTCGGCGTAGGCTACCGGTTTC 3201  
Db 3124 GTCGGCGTAGGCTACCGGTTTC 3144  
  
RESULT 16  
AAS43876  
ID AAS43876 standard; DNA; 4344 BP.  
XX AAS43876;  
XX AC  
XX AC  
DT 18-DEC-2001 (first entry)  
XX  
DE Neisseria meningitidis fusion protein delta-G983-961 DNA.  
XX  
KW Neisseria gonorrhoeae; leader peptide; fusion protein; ORF46.1; ds;  
Neisserial protein.  
XX  
OS Neisseria meningitidis.  
OS Synthetic.  
XX  
PN WO200164922-A2.  
XX  
PD 07-SEP-2001.  
XX  
PF 28-FEB-2001; 2001WO-1B00452.  
XX







Db 1684 AAGGAGTCTGCAGCTGGACGGCAAGGTACGCTGTACACAGCTTTGGGCAAACTGCTG 1743  
QY 1801 AAGTGSAGGTACGCGATGACCGCGCAAGCTGACATGTCGCGACGCGCAAGGG 1860  
Db 1744 AAGTGGAGCGGTACGCGGATATTCGGCGGCAAGCTGATACATGTCGCGACGCGCAAGGG 1803  
QY 1861 GCAGGCTATCTCAACCGTACCGGACAACTGTTCCCTCTCTAGTGCAGCGCAAAATCGGG 1920  
Db 1804 GCAGGCTATCTCAACAGTACCGGACGACGTGTTCCCTCTCTAGTGCAGCGCAAAATCGGG 1863  
QY 1921 CGGATATTTCTTCTTCAAAACATCGAAACGAGCGGTGCTGCTGGCTTCCCTCGAC 1980  
Db 1864 CAGGATATTTCTTCTTCAAAACATCGAAACGAGCGGCGGCTGCTGGCTTCCCTCGAC 1923  
QY 1981 AGCTCGAAAAAACAGCGGCGAGTGAAGCGGACACGCTGTCCTATTATGTCGTCGCGGC 2040  
Db 1924 AGCTCGAAAAAACAGCGGCGAGTGAAGCGGACACGCTGTCCTATTATGTCGTCGCGGC 1983  
QY 2041 AATGCGGCACGAGCTGCTTTCGGCAGCGGCACATTCGCGCGCCCGCTCTGAAACACGCG 2100  
Db 1984 AATGCGGCACGAGCTGCTTTCGGCAGCGGCACATTCGCGCGCCCGCTCTGAAACACGCG 2043  
QY 2101 GTAGAACAGGGCGGACGAATCTGGAACACCTGATGGTGCAGTGCATGCTCCGCAATCA 2160  
Db 2044 GTAGAACAGGGCGGACGAATCTGGAACACCTGATGGTGCAGTGCATGCTCCGCAATCA 2103  
QY 2161 TCGGCACACCCGAGACGCTTGAACATGCGCGCGCGCGACGCGACAGATATGCGCGGCATC 2220  
Db 2104 TCGGCACACCCGAGACGCTTGAACATGCGCGCGCGCGACGCGACAGATATGCGCGGCATC 2163  
QY 2221 CGCCCTACGCGGCAACTTTCGCGCAGCGGCAGCGGTACAGCATCGAATGCGCGCGAC 2280  
Db 2164 CGCCCTACGCGCAACTTTCGCGCAGCGGCAGCGGTACAGCATCGAATGCGCGCGAC 2223  
QY 2281 GGTGTACGATCTCAACAGTCTCGCGCTACCGTCTATCCGACAGTACCGCGCCCAT 2340  
Db 2224 GGTGTACGATCTCAACAGTCTCGCGCTACCGTCTATCCGACAGTACCGCGCCCAT 2283  
QY 2341 GCGGATATCAGGACGCGCGGTGAAGCGGTATCGAGCGGTGACACACAGCTACG 2400  
Db 2284 GCGGATATCAGGACGCGCGGTGAAGCGGTATCGAGCGGTGACACACAGCTACG 2343  
QY 2401 GGTCTCGCGCTCATCGCGCAAAACCCAGGACGCGGTGGAACGTTGGAAACAGGCGGTGT 2460  
Db 2344 GGTCTCGCGCTCATCGCGCAAAACCCAGGACGCGGTGGAACGTTGGAAACAGGCGGTGT 2403  
QY 2461 GAAGCAAAATGCGCGGAGTACCCAAACCGTCGCGCATTCGCGCAAAACCCGCGCAAAAT 2520  
Db 2404 GAAGCAAAATGCGCGGAGTACCCAAACCGTCGCGCATTCGCGCAAAACCCGCGCAAAAT 2463  
QY 2521 ACGACAGACGCGCCACACTGGCATGGGATGGGACAGCAGCATGGAGCGAAACAGTGC 2580  
Db 2464 ACGACAGACGCGCCACACTGGCATGGGATGGGACAGCAGCATGGAGCGAAACAGTGC 2523  
QY 2581 GCAAAAACCGACAGCATTAGTCTGTTTGCAGGATACGCGACGATCGCGCGATATCGGC 2640  
Db 2524 GCAAAACCGACAGCATTAGTCTGTTTGCAGGATACGCGACGATCGCGCGATATCGGC 2583  
QY 2641 TATCTCAAGGCGCTGTTCTCTACGACGCTACAAAACAGCATATCGCGCGACCGGT 2700  
Db 2584 TATCTCAAGGCGCTGTTCTCTACGACGCTACAAAACAGCATATCGCGCGACCGGT 2643  
QY 2701 GCGGACGACATCGGAGGACGCGTCAACGCGACCGCTATGACGCTGGCGGACATGGGC 2760  
Db 2644 GCGGACGACATCGGAGGACGCGTCAACGCGACCGCTATGACGCTGGCGGACATGGGC 2703  
QY 2761 GGTGTCAAGCTTCGCTTTCGCGCAAGGAGATTTTCACGCTCGAAGCGCTCTGCGCTAC 2820  
Db 2704 GGTGTCAAGCTTCGCTTTCGCGCAAGGAGATTTTCACGCTCGAAGCGCTCTGCGCTAC 2763  
QY 2821 GACCTGCTCAAAAGAGTGCATTCGCGCAAAAGGAGGAGTCTTTGGGCTGGAGCGCAAC 2880

Db 2764 GACCTGCTCAAAAGAGTGCATTCGCGCAAAAGGAGGAGTCTTTGGGCTGGAGCGCAAC 2823  
QY 2881 AGCTCAGTGAAGGACACACTGCTCGGCTCGGCTGCTGAAGCTGTCGCAACCTTTGAGC 2940  
Db 2824 AGCTCAGTGAAGGACACACTGCTCGGCTCGGCTGCTGAAGCTGTCGCAACCTTTGAGC 2883  
QY 2941 GATAAAGCCGCTCTCTGTTTGAACGCGGCGGTGAACGCGGCTGAACGCGGCTGAC 3000  
Db 2884 GATAAAGCCGCTCTCTGTTTGAACGCGGCGGTGAACGCGGCTGACGCGGCTGAC 2943  
QY 3001 ACGGTAAAGCGGCGCTTTTACCGGCGGCTGACGCAACCGGCAAGCGGCGGCTGAC 3060  
Db 2944 ACGGTAAAGCGGCGCTTTTACCGGCGGCTGACGCAACCGGCAAGCGGCGGCTGAC 3003  
QY 3061 ATCCCGCACACCGCTGTTTCCGCTGTCGGCGCGGATGTCGAATTCGGCAACGCTGG 3120  
Db 3004 ATCCCGCACACCGCTGTTTCCGCTGTCGGCGCGGATGTCGAATTCGGCAACGCTGG 3063  
QY 3121 AACGGCTTGGCAGCTTACAGCTACGCGGCTTCCAAACAGTACGCAACACGCGGACGA 3180  
Db 3064 AACGGCTTGGCAGCTTACAGCTACGCGGCTTCCAAACAGTACGCAACACGCGGACGA 3123  
QY 3181 GTCGCGTAGGCTACCGGCTTC 3201  
Db 3124 GTCGCGTAGGCTACCGGCTTC 3144

RESULT 17

AAS43874

ID AAS43874 standard; DNA: 4425 BP.

XX AAS43874;

XX DT 18-DEC-2001 (first entry)

XX Neisseria meningitidis fusion protein delta-G983-ORF46.1 DNA.

DE Neisseria gonorrhoeae; leader peptide; fusion protein; ORF46.1; ds;

KW Neisserial protein.

XX Neisseria meningitidis.

OS Synthetic.

XX WO200164922-A2.

XX PD 07-SEP-2001.

XX PF 28-FEB-2001; 2001WO-IB00452.

XX PR 28-FEB-2000; 2000GB-0004695.

XX PR 13-NOV-2000; 2000GB-0027675.

XX (CHIR-) CHIRON SPA.

PA Arico MB, Comanducci M, Galeotti C, Maignani V, Guilianini MM;

PI Pizsa M;

XX WPI; 2001-582163/65.

DR P-PSDB; AAU27575.

XX Producing heterologous proteins from Neisseria meningitidis and N.

XX gonorrhoeae -

XX Example 15; Page 42-43; 119pp; English.

CC The invention relates to methods for the heterologous expression of

CC Neisserial proteins from Neisseria meningitidis and Neisseria

CC gonorrhoeae. At least one domain in the protein is deleted, e.g. the

CC leader peptide, and may be replaced by a domain from a different protein

CC to make a fusion protein, in order to enhance heterologous expression of

CC Neisserial proteins. Also, a region of a protein, such as a poly-glycine

CC stretch, can be mutated to enhance expression. The proteins used in the

CC processes include ORF46.1, 287, 741, 919, 953, 961 and 983. Sequences

CC AAS43868-AAS43905 represent DNA molecules encoding Neisserial proteins  
XX and peptide regions of proteins of the invention.  
SQ Sequence 4425 BP; 1165 A; 1259 C; 1219 G; 782 T; 0 other;

Query Match 92.8%; Score 2971.8; DB 22; Length 4425;  
Best Local Similarity 97.2%; Pred. No. 0;  
Matches 3052; Conservative 0; Mismatches 47; Indels 42; Gaps 1;

Qy	103	ACTTCTGCGCCGACTCAATGTCAGCGCGCACCGGTATCGGCAGCAACAGACAGACACA	162
Db			
Qy	4	ACTTCTGCGCCGACTCAATGTCAGCGCGTACCGGTATCGGCAGCAACAGACAGACA	63
Db			
Qy	163	ACAGCGAAATCAGCAGCAGTATCTTACGCGGTATCAAGAAGCAAAATGTGCAAGACAGA	222
Db			
Qy	64	ACAGCGAAATCAGCAGCAGTATCTTACGCGGTATCAAGAAGCAAAATGTGCAAGACAGA	123
Db			
Qy	223	AGCATGCTGTGTCGGGTGGGATGACGTGCGGTTTACAGACAGGATGCCAAAATCAAT	282
Db			
Qy	124	AGCATGCTGTGTCGGGTGGGATGACGTGCGGTTTACAGACAGGATGCCAAAATCAAT	183
Db			
Qy	283	GCCTCCCTCCGAACTCGATACCGGACCTTTACAAACCCAAATGACGCATACAGAAAT	342
Db			
Qy	184	GCCTCCCTCCGAACTCGATACCGGACCTTTCCAAACCCAAATGACGCATACAGAAAT	243
Db			
Qy	343	TTGATCAACCTCAAACTTGAAGCAGGCTATACAGACGCGGGGTAGAGTAGGT	402
Db			
Qy	244	TTGATCAACCTCAAACTTGAAGCAGGCTATACAGACGCGGGGTAGAGTAGGT	303
Db			
Qy	403	ATCGTGTATACAGGCGAATTCGTCGGCAGCATATCCTTTCCGAACTGTATGGCAGAAA	462
Db			
Qy	304	ATCGTGTATACAGGCGAATTCGTCGGCAGCATATCCTTTCCGAACTGTATGGCAGAAA	363
Db			
Qy	463	GAACACGCTATAACGAAATTTACAAACTATACGCGTATATCGGAAGAACGCGCT	522
Db			
Qy	364	GAACACGCTATAACGAAATTTACAAACTATACGCGTATATCGGAAGAACGCGCT	423
Db			
Qy	523	GAAGACGAGCGGTAAAGACATTAAGCTTCTTTCGACGATGAGCGGTTATAGAGACT	582
Db			
Qy	424	GAAGACGAGCGGTAAAGACATTAAGCTTCTTTCGACGATGAGCGGTTATAGAGACT	483
Db			
Qy	583	GAAGCAAGCGACGGATATCGGCCACGTAAAGAAATCGGCACATCGATGTGCTCC	642
Db			
Qy	484	GAAGCAAGCGACGGATATCGGCCACGTAAAGAAATCGGCACATCGATTTGCTCC	543
Db			
Qy	643	CATATTATTGCGCGGTTCCGTGACGCGACCTGCGGCGGTATTGCGCCCGATCGG	702
Db			
Qy	544	CATATTATTGCGCGGTTCCGTGACGCGACCTGCGGCGGTATTGCGCCCGATCGG	603
Db			
Qy	703	ACGCTACACATAATGAATACGCATATGGAACCAAGAACGAAATATGTCTGCAGCCATC	762
Db			
Qy	604	ACGCTACACATAATGAATACGAATGATGAACCAAGAACGAAATGATGTTGCACCCATC	663
Db			
Qy	763	CGCAATGATGGTCAAGCTGGCGGAACGTGCGGTGCGCATCGTCAATAACAGTTTTGA	822
Db			
Qy	664	CGCAATGATGGTCAAGCTGGCGGAACGTGCGGTGCGCATCGTCAATAACAGTTTTGA	723
Db			
Qy	823	ACAACATCAGGGCAGCAGCTGCCACCATTTCCAAATAGCCAATTCGAGGAGCAGTAC	882
Db			
Qy	724	ACAACATCAGGGCAGCAGCTGCCACCATTTTCCAAAATAGCCAATTCGAGGAGCAGTAC	783
Db			
Qy	883	CGCCAAAGGTTGCTCGCCTATTTCGGCGGTGATAAAACAGACGAGGGTATCGCCTGATG	942
Db			
Qy	784	CGCCAAAGGTTGCTCGCCTATTTCGGCGGTGATAAAACAGACGAGGGTATCGCCTGATG	843
Db			
Qy	943	CAACAGAGGATTACGGCACTGTCTTACCACTATCCCGTAAATAAAGATGCTTTTCATT	1002
Db			
Qy	844	CAACAGAGGATTACGGCACTGTCTTACCACTATCCCGTAAATAAAGATGCTTTTCATC	903
Db			
Qy	1003	TTTTTCGCAAGCAATGACGCAAGCTCAGCCCAACACATGACCCCTATTGCCATTTTAT	1062
Db			
Qy	904	TTTTTCGCAAGCAATGACGCAAGCTCAGCCCAACACATATGCCCTATTGCCATTTTAT	963
Db			

Qy	1063	GAAAAAGATGCTCAAAAAGGCATTATATCAGTTCGAGCGGTAGACCGCAGTGGAGAAAG	1122
Db			
Qy	964	GAAAAAGAGCGTCAAAAAGGCATTATATCAGTTCGAGCGGTAGACCGCAGTGGAGAAAG	1023
Db			
Qy	1123	TTCA-----ATGCGTCCAACCAT	1140
Db			
Qy	1024	TTCAAACGGGAATGTATGGAAACCGGTTACAGAACCGCTTGAGTATGGTCTCAACCAT	1083
Db			
Qy	1141	TGCGGAATTTACTTGCCCATCTGTGTCCTATCGGCACCTATGAAGCAAGCGTCCGTTTCACC	1200
Db			
Qy	1084	TGCGGAATTTACTTGCCCATCTGTGTCCTGTGCGCACCTATGAAGCAAGCGTCCGTTTCACC	1143
Db			
Qy	1201	CGTACAAACCCGATTCAAATTTCCGGAACATCCTTTTCCGCACCCCATCGTAACCGGCAG	1260
Db			
Qy	1144	CGTACAAACCCGATTCAAATTTCCGGAACATCCTTTTCCGCACCCCATCGTAACCGGCAG	1203
Db			
Qy	1261	CGCGTCTGTGTCGACAGAAATACCCGTGGATGAGCAGCAGCAACCTCGGTACACAGCTG	1320
Db			
Qy	1204	CGCGTCTGTGTCGACAGAAATACCCGTGGATGAGCAACGCAACCTCGTACACAGCTG	1263
Db			
Qy	1321	CTGACAAACCGCTCAGGACATCGGTGACGTGCGGTGGACAGCAAGTTCCGTTGGGACTG	1380
Db			
Qy	1264	CTGACGACCGCTCAGGACATCGGTGACGTGCGGTGGACAGCAAGTTCCGTTGGGACTG	1323
Db			
Qy	1381	CTGATGCGGGTAAAGCCATGAACGACCCCGCTCTTTCCGTTCCGCGACTTTTACCGCC	1440
Db			
Qy	1324	CTGATGCGGGTAAAGCCATGAACGACCCCGCTCTTTCCGTTCCGCGACTTTTACCGCC	1383
Db			
Qy	1441	GATACGAAAGGTACATCCGATATTTCCCTACTCTTCCGTAAGCAGATTTTCAGGACGGGC	1500
Db			
Qy	1384	GATACGAAAGGTACATCCGATATTTCCCTACTCTTCCGTAAGCAGATTTTCAGGACGGGC	1443
Db			
Qy	1501	GGCTGTATCAAAAAGCGGACGCAACTGCAACTGACGCGCAACCAACCTATACGGGC	1560
Db			
Qy	1444	GGCTGTATCAAAAAGCGGACGCAACTGCAACTGACGCGCAACCAACCTATACGGGC	1503
Db			
Qy	1561	AAAACCATTTATCGAAGCGGTTGCTGTGTGTGTACGCAACCAACCAATCGATATCGC	1620
Db			
Qy	1504	AAAACCATTTATCGAAGCGGTTGCTGTGTGTGTACGCAACCAACCAATCGATATCGC	1563
Db			
Qy	1621	GTGCAAAACAAAGGTGCGCTGATTTTAAACGGGGGGGATCGGGGGTAGCCTGAACAGC	1680
Db			
Qy	1564	GTGCAAAACAAAGGTGCGCTGATTTTAAACGGGGGGGATCGGGGGTAGCCTGAACAGC	1623
Db			
Qy	1681	GACGCATTTGCTATCTGGCAGATACCGACCGCTCCGCGCAACCAACCTGTCACATC	1740
Db			
Qy	1624	GACGCATTTGCTATCTGGCAGATACCGACCGCTCCGCGCAACCAACCTGTCACATC	1683
Db			
Qy	1741	AAAGCGATCTGCACTGGCGGGGAAGGTACGCTGTACACACGTTTGGGCAAACTGCTG	1800
Db			
Qy	1684	AAAGCGATCTGCACTGGCGGGAAAGGTACGCTGTACACACGTTTGGGCAAACTGCTG	1743
Db			
Qy	1801	AAAGTGGAGGTACGGCGATGACCGGGGCAAGCTGTACATGTCCGGCAGCGGGCAAGGG	1860
Db			
Qy	1744	AAAGTGGAGGTACGGCGATTTATCGCGGCAAGCTGTACATGTCCGGCAGCGGGCAAGGG	1803
Db			
Qy	1861	GCAGCTATCTCAACCGTACCGGACACGTGTTCCCTTCCCTGAGTCCGCCAAAATCGGG	1920
Db			
Qy	1804	GCAGCTATCTCAACAGTACCGGACGAGTGTCCCTCTCTCTGAGTCCGCCAAAATCGGG	1863
Db			
Qy	1921	CGGGATTTCTTTCTTCCAAAACATCGAAACCGGAGTGTGCTGCTGGCTTCCCTCGAC	1980
Db			
Qy	1864	CAGGATTTATCTTTCTTCCAAAACATCGAAACCGGAGGCTGCTGGCTTCCCTCGAC	1923
Db			
Qy	1981	AGCTGCGAAAAACAGCGGGCAGTGAAGCGCACAGCTGTCTCTATTATGTCCGTCCGGC	2040
Db			
Qy	1924	AGCTGCGAAAAACAGCGGGCAGTGAAGCGCACAGCTGTCTCTATTATGTCCGTCCGGC	1983
Db			
Qy	2041	AATGGCGCACGACTGCTTTCCGACGGCACATTCGCGCGCGCGGCTGTGAACACAGCC	2100
Db			
Qy	1984	AATGGCGCACGACTGCTTTCCGACGGCACATTCGCGCGCGCGGCTGTGAACACAGCC	2043
Db			
Qy	2101	GTAGAACAGGGCGGACGCAATCTGGAAAAACCTGTGTTGGTGAACACTGGATGCCCTCCGAATCA	2160
Db			

Db	3124	GTCCGGCTAGGCTACCGGTTCC	3144
RESULT	18		
AAD17039			
ID	AAD17039	standard; DNA; 4425 BP.	
XX	AC		
XX	AAD17039;		
XX			
XX	29-NOV-2001	(first entry)	
XX			
DE	N. meningitidis strain 2996	delta G983-ORF46.1 fusion DNA.	
XX			
XX	Heterologous expression; Neisserial protein;	open reading frame; ORF; delta G983-ORF46.1 fusion protein; ds.	
KW			
KW			
XX	XX	Neisseria meningitidis 2996.	
OS			
XX			
FH	Key	Location/Qualifiers	
FT	CDS	1..4425	
FT		/*tag= a	
FT		/product= "N meningitidis strain 2996 delta	
FT		G983-ORF46.1 fusion protein"	
XX			
XX	W0200164920-A2.		
XX			
XX	07-SEP-2001.		
XX			
XX	28-FEB-2001; 2001WO-IB00420.		
PF			
XX			
XX	28-FEB-2000; 2000GB-0004695.		
PR			
PR	13-NOV-2000; 2000GB-0027675.		
XX			
XX	(CHIR-) CHIRON SPA.		
PA			
XX			
XX	Arico MB, Comanducci M, Galeotti C, Masignani V, Giuliani MM; Pizzo M;		
PI			
PT			
XX			
XX	WPI: 2001-557776/62.		
DR	P-PSDB; AAE10022.		
DR			
DR			
XX			
XX	Heterologous expression for the expression of two or more Neisserial proteins in fused state		
PT			
PT			
XX			
PS	Example 3; Page 13-14; 52pp; English.		
XX			
CC	The present invention relates to a method for simultaneous heterologous expression of two or more Neisserial proteins which are in a fused state. The method is useful for simultaneous heterologous expression of two or more Neisserial proteins. A protein that may be unstable or poorly expressed on its own is assisted by adding a suitable hybrid partner and commercial manufacture is simplified-only one expression purification need to be employed in order to produce two separately-useful proteins. The present sequence is a DNA encoding Neisseria meningitidis (serogroup B, strain 2996) delta G983-ORF46.1 (open reading frame) fusion protein.		
CC			
XX			
SQ	Sequence 4425 BP; 1165 A; 1259 C; 1219 G; 782 T; 0 other;		
Query Match	92.8%;	Score 2971.8;	DB 22; Length 4425;
Best Local Similarity	97.2%;	Pred. No. 0;	
Matches 3052;	Conservative	0; Mismatches	47; Indels 42; Gaps
Qy	103	ACTTCTCGGCCGCGACTTCAATGACGGGGACCGGGTATCGGCAGCAACAGCAGCAACA	162
	4	ACTTCTCGGCCGCGACTTCAATGACGGGGTATCGGCAGCAACAGCAGCAACA	63
Db			
Qy	163	ACACGAAATCAGCAGCAGTATCTTAGCCCGGTATCAAGAACAAATGTCCAAGACAGA	222
Db	64	ACACGAAATCAGCAGCAGTATCTTAGCCCGGTATCAAGAACAAATGTCCAAGACAGA	123
Qy	223	AGCATGCTCTGTGCGGTCGGGATGAGTTGCGGTTACAGACAGGATGCCAAATCAAT	282

Db 124 AGCATGCTGTGCGGGTGGGATGACGTTGCGGTTTACAGACAGGATGCCAAAATCAAT. 183  
Qy 283 GCCCCCCCGGAATCTGCATACCGGAGACTTTACAAACCCAAATGACGATACAGAAT. 342  
Db 184 GCCCCCCCGGAATCTGCATACCGGAGACTTTCCAAACCCAAATGACGATACAGAAT. 243  
Qy 343 TTGATCAACCTCAAAACCTGCAATTTGAAGCAGGCTATACAGACGCGGGGTAGAGTAGT. 402  
Db 244 TTGATCAACCTCAAAACCTGCAATTTGAAGCAGGCTATACAGACGCGGGGTAGAGTAGT. 303  
Qy 403 ATCTCGATACAGCGAATCCGTGCGCAGCATATCCTTTCCGAACTGTATGGCAGAAA. 462  
Db 304 ATCTCGACACAGCGAATCCGTGCGCAGCATATCCTTTCCGAACTGTATGGCAGAAA. 363  
Qy 463 GAACACGGCTATAACGAAATTTACAAAACTATACGGCGTATATCGCGAAGGAAAGCGCT. 522  
Db 364 GAACACGGCTATAACGAAATTTACAAAACTATACGGCGTATATCGCGAAGGAAAGCGCT. 423  
Qy 523 GAAGACGGAGCGGTAAAGCATTAAAGCTTCTTTTCGACGATAGGCGGTTTATAGAGACT. 582  
Db 424 GAAGACGGAGCGGTAAAGCATTAAAGCTTCTTTTCGACGATAGGCGGTTTATAGAGACT. 483  
Qy 583 GAAGCAAGCGCGGATATCCGCCAGCTAAAGAAATCGACACATCGATGTGCTCTCC. 642  
Db 484 GAAGCAAGCGCGGATATCCGCCAGCTAAAGAAATCGACACATCGATGTGCTCTCC. 543  
Qy 643 CATATTATTGGCGGCTTCCGTGGACGGCAGACTCGCAGCGGTATTGCGCCGATGCG. 702  
Db 544 CATATTATTGGCGGCTTCCGTGGACGGCAGACTCGCAGCGGTATTGCGCCGATGCG. 603  
Qy 703 ACGCTACACATAATGAAATACGATGATGGAACCAAGAAATGATGCTGCGACGATC. 762  
Db 604 ACGCTACACATAATGAAATACGATGATGGAACCAAGAAATGATGCTGCGACGATC. 663  
Qy 763 CGCAATCGATGGTCAAGCTGGCGGAAGCTGGCGTCCGATCGATCAATAACAGTTTGA. 822  
Db 664 CGCAATCGATGGTCAAGCTGGCGGAAGCTGGCGTCCGATCGATCAATAACAGTTTGA. 723  
Qy 823 ACAACATCGAGGCGGACACTGCGGACCATTTCCAAATAGCAATTCGGAGGACGATAC. 882  
Db 724 ACAACATCGAGGCGGACACTGCGGACCATTTCCAAATAGCAATTCGGAGGACGATAC. 783  
Qy 883 CGCAAGGGTGTGCGCTATTCCGGGGTGATAAACAGACGAGGATCCGCGCTGATG. 942  
Db 784 CGCAAGGGTGTGCGCTATTCCGGGGTGATAAACAGACGAGGATCCGCGCTGATG. 843  
Qy 943 CAACAGAGGATACGGCAACTTGTCTACCATCCGATCCGTAATAAAACATGCTTTTCAT. 1002  
Db 844 CAACAGAGGATACGGCAACTTGTCTACCATCCGATCCGTAATAAAACATGCTTTTCATC. 903  
Qy 1003 TTTTCGGCAAGCAATGACGCACAAGCTCAGCCCAACACTGACCTATTGCCATTTTAT. 1062  
Db 904 TTTTCGACAGCAATGACGCACAAGCTCAGCCCAACACTGATGCCATTTTAT. 963  
Qy 1063 GAAAAAGATGCTCAAAAGGCAATATACAGTCGACGCGTAGACCGAGTGGAGAAAG. 1122  
Db 964 GAAAAAGATGCTCAAAAGGCAATATACAGTCGACGCGTAGACCGAGTGGAGAAAG. 1023  
Qy 1123 TTCA-----ATGGCTCCAACCAT. 1140  
Db 1024 TTCAACGGGAAATGTATGGAGAACCGGGTACAGAACCGCTTGAGATATGGCTCCAACCAT. 1083  
Qy 1141 TCGGAATTTACTGCGCATGTGGCTATPCGGCACCCTATGAGCAAGCGTCCGTTTACC. 1200  
Db 1084 TCGGAATTTACTGCGCATGTGGCTGTCCGACCCCTATGAGCAAGCGTCCGTTTACC. 1143  
Qy 1201 CGTCAAAACCGGATTCAAATTTCCGGGAACATCTCTTTCCGCCACCCATCGTAACCGGCACG. 1260  
Db 1144 CGTCAAAACCGGATTCAAATTTCCGGGAACATCTCTTTCCGCCACCCATCGTAACCGGCACG. 1203  
Qy 1261 CGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACGATCCGTAACACGCTG. 1320  
Db 1204 CGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACGATCCGTAACACGCTG. 1263

Qy 1321 CTGACAAACGGCTCAGGACATCGGTGCGTGCAGTCGGGTGACAGCAAGTTTCGGTGGGACTG. 1380  
Db 1264 CTGACGAGCGCTCAGGACATCGGTGCGTGCAGTCGGGTGACAGCAAGTTTCGGTGGGACTG. 1323  
Qy 1381 CTGATCGGGTAAAGCCATGAAAGGACCGCGCTCTTTCCGTTCCGGCTGAGCTTTACGCC. 1440  
Db 1324 CTGATCGGGTAAAGCCATGAAAGGACCGCGCTCTTTCCGTTCCGGCTGAGCTTTACGCC. 1383  
Qy 1441 GATACGAAAGTACATCCGATATTGCTACTCTCTTCGTTAACGACATTTTCAGGACCGGC. 1500  
Db 1384 GATACGAAAGTACATCCGATATTGCTACTCTCTTCGTTAACGACATTTTCAGGACCGGC. 1443  
Qy 1501 GGCCTGATCAAAAAGCGGCGCAACTGCAACTGCACGSCAAACACCTATACGGC. 1560  
Db 1444 GGCCTGATCAAAAAGCGGCGCAACTGCAACTGCACGSCAAACACCTATACGGC. 1503  
Qy 1561 AAAACCAATTATCGAAGCGGTTCTGCTGTTGTACGGCAACAAATCGGATATGCG. 1620  
Db 1504 AAAACCAATTATCGAAGCGGTTCTGCTGTTGTACGGCAACAAATCGGATATGCG. 1563  
Qy 1621 GTCGAAACCAAGTGCCTGATTTATAACGGGCGCATCCGCGTAGCTCTGAACAGC. 1680  
Db 1564 GTCGAAACCAAGTGCCTGATTTATAACGGGCGCATCCGCGTAGCTCTGAACAGC. 1623  
Qy 1681 GACGGCATTTCTATCTGGCAGATACCGACCGATCCGGCGCAACACGCTGACATC. 1740  
Db 1624 GACGGCATTTCTATCTGGCAGATACCGACCAATCCGCGCAACACGCTGACATC. 1683  
Qy 1741 AAGGCGATCTGACGTGGGCGGAAGGTACGCTGTACACACGTTTGGGAAACTGCTG. 1800  
Db 1684 AAGGCGATCTGACGTGGGCGGAAGGTACGCTGTACACACGTTTGGGAAACTGCTG. 1743  
Qy 1801 AAGTGACGCTACGGGATGACCGCGCAAGCTGTACATGTCGGCACGCGCAAGAGG. 1860  
Db 1744 AAGTGACGCTACGGGATGATCGCGCAAGCTGTACATGTCGGCACGCGCAAGAGG. 1803  
Qy 1861 GAGGCTATCTCAACCGTACCGGACACGTTTCCCTTCTCTGAGTCGCGCAAAATCGG. 1920  
Db 1804 GAGGCTATCTCAACAGTACCGGACACGTTTCCCTTCTCTGAGTCGCGCAAAATCGG. 1863  
Qy 1921 CGGATTTATCTTTTTCACAAACATCGAAACCGAGCGTGTCTTATTTGCTCGCGGC. 1980  
Db 1864 CAGATTTATCTTTTTCACAAACATCGAAACCGAGCGCTGCTGCTGCTGCTGCTG. 1923  
Qy 1981 AGCGTCAAAAACAGCGGCGAGTGAAGGCGACACGCTGCTTATTTGCTCGCGGC. 2040  
Db 1924 AGCGTCAAAAACAGCGGCGAGTGAAGGCGACACGCTGCTTATTTGCTCGCGGC. 1983  
Qy 2041 AATCGGCGAGCTGCTTTCGGCAGCGCACATTTCCGCGCGCGCGCTCTGAAACACGCC. 2100  
Db 1984 AATCGGCGAGCTGCTTTCGGCAGCGCACATTTCCGCGCGCGCGCTCTGAAACACGCC. 2043  
Qy 2101 GTAGAACAGGGCGGCGCAATCTCGAANAACCTGATGTGAACTGGATGCTCTCGCAATCA. 2160  
Db 2044 GTAGAACAGGGCGGCGCAATCTCGAANAACCTGATGTGAACTGGATGCTCTCGCAATCA. 2103  
Qy 2161 TCCGCAACCCCGAGAGCGTTGAACTGCGCGCGCGCACGACAGATATGCGCGGCATC. 2220  
Db 2104 TCCGCAACCCCGAGAGCGTTGAACTGCGCGCGCGCACGACAGATATGCGCGGCATC. 2163  
Qy 2221 CGCCCTACGGCGCAACTTTTCGGCGCAGCGCACCGCTTACAGCATGGAATGCGCGGCAC. 2280  
Db 2164 CGCCCTACGGCGCAACTTTTCGGCGCAGCGCACCGCTTACAGCATGGAATGCGCGGCAC. 2223  
Qy 2281 GGTGTAGCATCTTCAACAGTCTCGCGCTACCGCTCTATGCCGACAGTACCGCGCCAT. 2340  
Db 2224 GGTGTAGCATCTTCAACAGTCTCGCGCTACCGCTCTATGCCGACAGTACCGCGCCAT. 2283  
Qy 2341 GCCGATGAGGACCGCGCTGAAAGCGTATCGACGGGTTGGACCAACAGCCTACG. 2400  
Db 2284 GCCGATGAGGACCGCGCTGAAAGCGTATCGACGGGTTGGACCAACAGCCTACG. 2343



```
QY 463 GAACACGGCTATAACGAAATTTACAAAACATATACGGCGTATATCCGGAAGAAACGGCT 522
    |||||
Db 364 GAACACGGCTATAACGAAATTTACAAAACATATACGGCGTATATCCGGAAGAAACGGCT 423
QY 523 GAAGACGAGGCGGTAAGACATTAAGCTTCTTCGACGATGAGCGCTTATAGAGACT 582
    |||||
Db 424 GAAGACGAGGCGGTAAGACATTAAGCTTCTTCGACGATGAGCGCTTATAGAGACT 483
QY 583 GAAGCAAAACCGACGATATCCGCCACCTFAAAGAAATCGGACACATCATGTGTCTCC 642
    |||||
Db 484 GAAGCAAAACCGACGATATCCGCCACCTFAAAGAAATCGGACACATCATGTGTCTCC 543
QY 643 CATATTATTGGCGGCTTCGCTGACGCGACACTGACGCGGTATTGCGCCCGATGCG 702
    |||||
Db 544 CATATTATTGGCGGCTTCGCTGACGCGACACTGACGCGGTATTGCGCCCGATGCG 603
QY 703 AGCTTACACATTAATGAATACGATGATGAACCAAGACGAAATTAATGCTGCAACCATC 762
    |||||
Db 604 AGCTTACACATTAATGAATACGATGATGAACCAAGACGAAATTAATGCTGCAACCATC 663
QY 763 CGCAATGATGGTCAAGCTGGCGAAGCTGGCGTGCSCATCGTCAATAAAGATTTTGA 822
    |||||
Db 664 CGCAATGATGGTCAAGCTGGCGAAGCTGGCGTGCSCATCGTCAATAAAGATTTTGA 723
QY 823 ACAACATCGAGGCGAGGCACTGCGGACCATTTCCAAATAGCCAAATTCGGAGGAGCAGTAC 882
    |||||
Db 724 ACAACATCGAGGCGAGGCACTGCGGACCATTTTCCAAATAGCCAAATTCGGAGGAGCAGTAC 783
QY 883 CGCCAAAGCTTGCTCGCTATTCGCGCGGTGATAAAACAGACGAGGGTATCGGCGTGATG 942
    |||||
Db 784 CGCCAAAGCTTGCTCGCTATTCGCGCGGTGATAAAACAGACGAGGGTATCGGCGTGATG 843
QY 943 CAACAGAGCGATTACGGCAACTTGTCTACACATCCGTAATAAAACATGCTTTTCATT 1002
    |||||
Db 844 CAACAGAGCGATTACGGCAACTTGTCTACACATCCGTAATAAAACATGCTTTTCATT 903
QY 1003 TTTTCGCGCAAGCAATGAGCACAAGCTCAGCCCCACACACTGACCCCTATTGCGCAATTTAT 1062
    |||||
Db 904 TTTTCGCGCAAGCAATGAGCACAAGCTCAGCCCCACACACTGACCCCTATTGCGCAATTTAT 963
QY 1063 GAAAAGATGCTCAAAAGGCAATATACAGTCGCGAGGGTAGCGGAGTGAGGGAAG 1122
    |||||
Db 964 GAAAAGAGCGCTCAAAAGGCAATATACAGTCGCGAGCGTAGACCGCAGTGGAGAAAG 1023
QY 1123 TTCA-----ATGGCTCCAACCAT 1140
    |||||
Db 1024 TTCAAGCGGAAATGATGAGAGAACGGGTACAGAACCGCTTGAGTATGGCTCCAACCAT 1083
QY 1141 TCGGAATTTACTGCCATGTGTSCCTATCGGCACCCCTATGAAGCAAGCGTCCGTTTCACC 1200
    |||||
Db 1084 TCGGAATTTACTGCCATGTGTSCCTATCGGCACCCCTATGAAGCAAGCGTCCGTTTCACC 1143
QY 1201 CGTACAAACCCGATTCAAATTCGCGGACATCCTTTTCGGCACCCATCGTAACCGGCAGG 1260
    |||||
Db 1144 CGTACAAACCCGATTCAAATTCGCGGAAACATCCTTTTCGGCACCCATCGTAACCGGCAGG 1203
QY 1261 GCGGCTCTGCTCTGCAGAAATACCGGTGATGAGCAACGACAACTCGGTACACAGCTG 1320
    |||||
Db 1204 GCGGCTCTGCTCTGCAGAAATACCGGTGATGAGCAACGACAACTCGGTACACAGCTG 1263
QY 1321 CTGACAACGGCTCAGGACATCGGTGACGTCGCGGTGGACAGCAAGTTCGGCTGGGCACTG 1380
    |||||
Db 1264 CTGACGACGGCTCAGGACATCGGTGACGTCGCGGTGGACAGCAAGTTCGGCTGGGCACTG 1323
QY 1381 CTGGATGCGGGTAAGCCCATGAAAGGACCGCGGTCTTCGCTTCGGGACTTTACCGGC 1440
    |||||
Db 1324 CTGGATGCGGGTAAGCCCATGAAAGGACCGCGGTCTTCGCTTCGGGACTTTACCGGC 1383
QY 1441 GATACGAAAGGTACATTCGATATTCCTTCTTCCGTTCAGGACATTTTCAGGACGCGG 1500
    |||||
Db 1384 GATACGAAAGGTACATTCGATATTCCTTCTTCCGTTCAGGACATTTTCAGGACGCGG 1443
QY 1501 GGCCTGATCAAAAAAGGCGGACCGCAACTGCAACTGCACGGCAACAACACCTTATACGGGC 1560
```

```
Db 1444 GGCCTGATCAAAAAAGGCGGACCGCAACTGCAACTGCAAGCAACAACACCTATACGGGC 1503
    |||||
QY 1561 AAAACCATATCGAAGCGGTTGCTGTGTGTGTACGCAACAACAATCGGATGCGC 1620
    |||||
Db 1504 AAAACCATATCGAAGCGGTTGCTGTGTGTGTACGCAACAACAATCGGATGCGC 1563
    |||||
QY 1621 GTCGAAACCAAAAGGTGCGCTGATTATTAACGGGGCGGATCCGGCGTAGCTGAACAGC 1680
    |||||
Db 1564 GTCGAAACCAAAAGGTGCGCTGATTATTAACGGGGCGGATCCGGGGGAGCTGNACAGC 1623
    |||||
QY 1681 GACGCAATTGTCTATCTGCGCAGATACCGACCGATCCGCGCAAAAGCAACCGTCGCATTC 1740
    |||||
Db 1624 GACGCAATTGTCTATCTGCGCAGATACCGACCAATCCGCGCAAAAGCAACCGTCACATC 1683
    |||||
QY 1741 AAAGCGCATCTGCAGCTGGGCGGGAAGGTACGCTGTACACACGTTTGGGCAAACTGCTG 1800
    |||||
Db 1684 AAAGCGCATCTGCAGCTGGGCGGCAAGGTACGCTGTACACACGTTTGGGCAAACTGCTG 1743
    |||||
QY 1801 AAAGTGGACGCTAGCGGCGATGACCGGGCGCAAGCTGTACATGTTCGGCACGCGGCAAGGG 1860
    |||||
Db 1744 AAAGTGGACGCTAGCGGCGATTATCGGCGGCAAGCTGTACATGTTCGGCACGCGGCAAGGG 1803
    |||||
QY 1861 GCAGCTATCTCAACCGGTACCGGACACAGTGTTCCTTCTGTAGTGCAGCAAAATCGGG 1920
    |||||
Db 1804 GCAGCTATCTCAACAGTACCGGACAGCTGTTCCTTCTGTAGTGCAGCAAAATCGGG 1863
    |||||
QY 1921 CGGATTTATCTTCTTCCAAAACATCGAAAACGAGGTGTCTGCTGGCTTCCCTCCAC 1980
    |||||
Db 1864 CAGGATTTATCTTCTTCCAAAACATCGAAAACGAGGTGTCTGCTGGCTTCCCTCCAC 1923
    |||||
QY 1981 AGCGTCGAAAAAAGCGGGCAGTGAAGCGCACACGCTGTCTATTATGTCCGTCCGCGC 2040
    |||||
Db 1924 AGCGTCGAAAAAAGCGGGCAGTGAAGCGCACACGCTGTCTATTATGTCCGTCCGCGC 1983
    |||||
QY 2041 AATCGGCGACGACTGCTTCGGCAGCGCACATTCGCGCGCGCGGGTCTGAAACACGCC 2100
    |||||
Db 1984 AATCGGCGACGACTGCTTCGGCAGCGCACATTCGCGCGCGCGGGTCTGAAACACGCC 2043
    |||||
QY 2101 GTAGAACAGGCGGCGAGCAATCTGAAAAACCTGATGGTTCGAACTGGATCCCTCCGAATCA 2160
    |||||
Db 2044 GTAGAACAGGCGGCGAGCAATCTGAAAAACCTGATGGTTCGAACTGGATCCCTCCGAATCA 2103
    |||||
QY 2161 TCC 2163
    |||
Db 2104 TCC 2106

RESULT 20
AAZ53301
ID AAZ53301 standard; DNA; 1365 BP.
XX
AC AAZ53301;
XX
DT 21-MAR-2000 (first entry)
XX
DE Neisseria meningitidis ORF 140 partial DNA sequence SEQ ID NO:551.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW antibacterial; gene therapy; ds.
XX
OS Neisseria meningitidis.
XX
PN WO9957280-A2.
XX
PD 11-NOV-1999.
XX
PF 30-APR-1999; 99WO-US09346.
XX
PR 01-MAY-1998; 98US-0083758.
PR 31-JUL-1998; 98US-0094869.
PR 02-SEP-1998; 98US-0098994.
```

PR 02-SEP-1998; 98US-0099062.  
PR 09-OCT-1998; 98US-0103749.  
PR 09-OCT-1998; 98US-0103794.  
PR 09-OCT-1998; 98US-0103796.  
PR 25-FEB-1999; 98US-0121528.  
XX (CHIR ) CHIRON CORP.  
PA (GENO-) INST GENOMIC RES.  
XX  
XX Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;  
PI Petersen J, Pizsa M, Rappuoli R, Ratti G, Scalato E, Scarselli M;  
PI Tettelin H, Venter JC;  
XX  
DR WPI: 2000-062150/05.  
DR P-PSDB; AAU74539.  
XX  
PT Novel Neisserial polypeptides predicted to be useful antigens for  
PT vaccines and diagnostics -  
XX  
PS Claim 7; Page 396-397; 1453pp; English.  
XX  
XX AAU53015 to AAU54536, AAU54577 to AAU54615, and AAU74253 to AAU75941  
CC represent novel *Neisseria meningitidis* and *N. gonorrhoeae* polynucleotides  
CC and polypeptides. AAU54537 to AAU54576 and AAU54616 to AAU55473 represent  
CC PCR primers used in the exemplification of the present invention. The  
CC polypeptides, the polynucleotides, antibodies and compositions of  
CC the invention can be used as vaccines, as diagnostic reagents, and as  
CC immunogenic compositions. The polypeptides can be used in the  
CC manufacture of medicaments for treating or preventing infection due to  
CC *Neisseria* bacteria (e.g. *Neisseria meningitidis* and *Neisseria gonorrhoeae*), to detect the  
CC presence of *Neisseria* bacteria, or to raise antibodies. They may also  
CC be used to screen for agonists or antagonists, which may themselves  
CC have use as antibacterial agents. The polynucleotides of the invention  
CC may also be used in gene therapy protocols.  
XX  
XX Sequence 1365 BP; 319 A; 410 C; 418 G; 218 T; 0 other;  
XX  
Query Match 42.1%; Score 1349; DB 21; Length 1365;  
Best Local Similarity 99.3%; Pred. No. 0;  
Matches 1355; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
QY 1840 ATGTGGCAGCGGCAAGGGGAGGGTATCTCAACGCTACCGGACAACTGTTCCTTC 1899  
DB 1 ATGTGGCAGCGGTAAGGGGAGGGTATCTCAACGCTACCGGACAACTGTTCCTTC 60  
QY 1900 CTGAGTCCGCCAAATCGGGGGGATTTATCTTCTCAACAACTCGAAACCGAGGT 1959  
DB 61 CTGAGTCCGCCAAATCGGGGGGATTTATCTTCTCAACAACTCGAAACCGAGCG 120  
QY 1960 GGCTCTGCTTCCCTCGACAGCGTGGAAAAACAGCGGGGAGTGAAGCGACAGCGTG 2019  
DB 121 GGCTCTGCTTCCCTCGACAGCGTGGAAAAACAGCGGGGAGTGAAGCGACAGCGTG 180  
QY 2020 TCTATATATGTCGTCGCGCAATGGGCGAGCGTCTTCGCGAGGGGACATTCGCGG 2079  
DB 181 TCTATATATGTCGTCGCGCAATGGGCGAGCGTCTTCGCGAGGGGACATTCGCGG 240  
QY 2080 CCGCGCGGTCTGAACACAGCGGTAGAACAGCGGCGGAGCAATCTGGAACACCTGATGGTC 2139  
DB 241 CCGCGCGGTCTGAACACAGCGGTAGAACAGCGGCGGAGCAATCTGGAACACCTGATGGTC 300  
QY 2140 GAATGGATGCTTCCGAATCATCCGAACACCGCGAGAGCGGTGGAACCTGGCGCGCGAC 2199  
DB 301 GAATGGATGCTTCCGAATCATCCGAACACCGCGAGAGCGGTGGAACCTGGCGCGCGAC 360  
QY 2200 CGCACAGATATCGGGGATCCGCCCTACGGCGCACTTTTCGCGAGGGGAGCGCGTA 2259  
DB 361 CGCACAGATATCGGGGATCCGCCCTACGGCGCACTTTTCGCGAGGGGAGCGCGTA 420  
QY 2260 CAGCATGCGAATCCGCCGAGCGGTAGCATCTTCAACAGTCTCGCGCTACCGTCTAT 2319  
DB 421 CAGCATGCGAATCCGCCGAGCGGTAGCATCTTCAACAGTCTCGCGCTACCGTCTAT 480

QY 2320 GCCGACAGTACCGCGCCCATGCCATATGACGGAGCGCGGTGAAAGCCGTATCGGAC 2379  
DB 481 GCCGACAGTACCGCGCCCATGCCATATGACGGAGCGCGGTGAAAGCCGTATCGGAC 540  
QY 2380 GGGTTGACACAAACGCTACGGGTCTCGCGCTCATCGCGAAACCCAAACAGGAGCGTGA 2439  
DB 541 GGGTTGACACAAACGCTACGGGTCTCGCGCTCATCGCGAAACCCAAACAGGAGCGTGA 600  
QY 2440 ACCTGGGAACAGCGCGGTGTGAAGGCAAAATGCGCGGAGTACCCAAACCCGTGCGCAT 2499  
DB 601 ACCTGGGAACAGCGCGGTGTGAAGGCAAAATGCGCGGAGTACCCAAACCCGTGCGCAT 660  
QY 2500 GCCCGGAAACCGCGGCAAAATGACAGAGCGCGCACACTGGGCAATGGGACACAGCACA 2559  
DB 661 GCCCGGAAACCGCGGCAAAATGACAGAGCGCGCACACTGGGCAATGGGACACAGCACA 720  
QY 2560 TGGAGCGAAACAGTGCATAATGCAAAACCGACAGCACTTGTGTGTTGAGGAGTACCGG 2619  
DB 721 TGGAGCGAAACAGTGCATAATGCAAAACCGACAGCACTTGTGTGTTGAGGAGTACCGG 780  
QY 2620 CAGGATCGCGGCGATATCGGCTATCTCAAGGGCTGTCTCTACGAGCGCTACAAAAAC 2679  
DB 781 CAGGATCGCGGCGATATCGGCTATCTCAAGGGCTGTCTCTACGAGCGCTACAAAAAC 840  
QY 2680 AGCATCAGCGCAGCAGCGGTGCGGAGCAATGCGGAAGCGAGCGTCAACGCGACGCTG 2739  
DB 841 AGCATCAGCGCAGCAGCGGTGCGGAGCAATGCGGAAGCGAGCGTCAACGCGACGCTG 900  
QY 2740 ATCAGTGGGCGACTGGCGGTGTCAACGTTCCGTTTCCGCAACGGGAGATTTGACG 2799  
DB 901 ATCAGTGGGCGACTGGCGGTGTCAACGTTCCGTTTCCGCAACGGGAGATTTGACG 960  
QY 2800 GTCAAGGGCGTCTCGCTACGAGCTCTCAACAGGATCATTCGCCGAAAGGAGCAGT 2859  
DB 961 GTCAAGGGCGTCTCGCTACGAGCTCTCAACAGGATCATTCGCCGAAAGGAGCAGT 1020  
QY 2860 GCTTTGGGCTGGAGCGCAACAGCGCTCACTGAAGGACACACTGGTTCGAGCTCGCGGGTCTG 2919  
DB 1021 GCTTTGGGCTGGAGCGCAACAGCATCACTGAAGGACACACTGGTTCGAGCTCGCGGGTCTG 1080  
QY 2920 AAGCTGTGCAACCCCTTGAGCGATAAAGCGTCTCTGTTTGCACAGCGGCGGTGGAACGC 2979  
DB 1081 AAGCTGTGCAACCCCTTGAGCGATAAAGCGTCTCTGTTTGCACAGCGGCGGTGGAACGC 1140  
QY 2980 GACCTGAACGCGCGACTACACGTTAACGGCGGCTTTTACCGCGCGACTGCGCAAC 3039  
DB 1141 GACCTGAACGCGCGACTACACGTTAACGGCGGCTTTTACCGCGCGACTGCGCAAC 1200  
QY 3040 GGCAAGACGGGCGCAGCAATATCGCGCACACCGCGCTGTTGCGCGGTCTGCGCGGGAT 3099  
DB 1201 GGCAAGACGGGCGCAGCAATATCGCGCACACCGCGCTGTTGCGCGGTCTGCGCGGGAT 1260  
QY 3100 GTCAATTCGGCAACGGCTGGAGCGGTGGACAGTTTACAGTACGCGGTTCACAAACAG 3159  
DB 1261 GTCAATTCGGCAACGGCTGGAGCGGTGGACAGTTTACAGTACGCGGTTCACAAACAG 1320  
QY 3160 TACGGCAACACAGCGAGCGAGTGGGAGTACCGGTTCCTGA 3204  
DB 1321 TACGGCAACACAGCGAGCGAGTGGGAGTACCGGTTCCTGA 1365  
RESULT 21  
AAU53300  
ID AAU53300 standard; DNA; 1365 BP.  
XX  
AC AAU53300;  
XX  
DT 21-MAR-2000 (first entry)  
XX  
DE *Neisseria meningitidis* ORF 140 partial DNA sequence SEQ ID NO:549.  
XX  
KW *Neisseria meningitidis*; *Neisseria gonorrhoeae*; antigen; vaccine;  
antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;



antibacterial; gene therapy; ds.

Neisseria meningitidis.  
 WO957280-A2.  
 11-NOV-1999.  
 30-APR-1999; 99WO-US09346.  
 01-MAY-1998; 98US-0083758.  
 31-JUL-1998; 98US-0094869.  
 02-SEP-1998; 98US-0098994.  
 02-SEP-1998; 98US-0099062.  
 09-OCT-1998; 98US-0103749.  
 09-OCT-1998; 98US-0103794.  
 09-OCT-1998; 98US-0103796.  
 25-FEB-1999; 99US-0121528.  
 (CHIR ) CHIRON CORP.  
 PA (GENO-) INST GENOMIC RES.  
 XX  
 PI Fraser C, Galeotti C, Grandi G, Hickey E, Massignani V, Mora M;  
 PI Petersen J, Pizsa M, Rappuoli R, Ratti G, Scalato E, Scarselli M;  
 PI Tettelin H, Venter JC;  
 XX  
 DR WPI: 2000-062150/05.  
 DR P-PSDB; AAY74538.  
 XX  
 Novel Neisserial polypeptides predicted to be useful antigens for  
 PT vaccines and diagnostics -  
 XX  
 XX Claim 7; Page 395; 1453pp; English.  
 PS  
 CC AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941  
 . CC represent novel Neisseria meningitis and N. gonorrhoea polynucleotides  
 CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent  
 CC PCR primers used in the exemplification of the present invention. The  
 CC polypeptides, the polynucleotides, antibodies and compositions of  
 CC the invention can be used as vaccines, as diagnostic reagents, and as  
 CC immunogenic compositions. The polypeptides can be used in the  
 CC manufacture of medicaments for treating or preventing infection due to  
 CC Neisserial bacteria (e.g. meningitis and septicaemia), to detect the  
 CC presence of Neisseria bacteria, or to raise antibodies. They may also  
 CC be used to screen for agonists or antagonists, which may themselves  
 CC have use as antibacterial agents. The polynucleotides of the invention  
 CC may also be used in gene therapy protocols.  
 XX  
 SQ Sequence 1365 BP; 317 A; 413 C; 421 G; 214 T; 0 other;  
 Query Match 41.9%; Score 1342.6; DB 21; Length 1365;  
 Best Local Similarity 99.0%; Pred. No. 0;  
 Matches 1351; Conservative 0; Mismatches 14; Indels 0; Gaps 0;  
 QY 1840 ATGTCGGCAGCGGCAAGGGGAGGCTATCTCAACCGTACCGGACACAGTGTTCCTTC 1899  
 DB 1 ATCTCGGCACGCGGCAAGGGGAGGCTATCTCAACAGTACCGGACGAGTGTTCCTTC 60  
 QY 1900 CTGAGTCCCGCCAAATCGGGCGGATTATCTTTCTTCAACAACATCGAAACCGAGGT 1959  
 DB 61 CTGAGTCCCGCCAAATCGGGCGGATTATCTTTCTTCAACAACATCGAAACCGAGCG 120  
 QY 1960 GGTCTGCTGGCTTCCCTCGACAGCGTCAAAAAAAGCGGGAGTGAAGCGACACGCTG 2019  
 DB 121 GGCTGCTGGCTTCCCTCGACAGCGTCAAAAAAAGCGGGAGTGAAGCGACACGCTG 180  
 QY 2020 TCCTATTATGTCGTCGCGCAATCGGGCAGCGACTGCTTCGGCAGCGGCACATCCCGG 2079  
 DB 181 TCCTATTATGTCGTCGCGCAATCGGGCAGCGACTGCTTCGGCAGCGGCACATCCCGG 240  
 QY 2080 CCGCGCGGTCTGAAACACGCGGTAGAACAGGGCGGAGCAATCTGAAACCTGATGGTC 2139  
 DB 241 CCGCGCGGTCTGAAACACGCGGTAGAACAGGGCGGAGCAATCTGAAACCTGATGGTC 300

QY 2140 GAACTGGATGCTCCGAATCATCCGCAACACCCGAGAGCGTTGAAACTGCGCCCGCGAC 2199  
 DB 301 GAACTGGATGCTCCGAATCATCCGCAACACCCGAGAGCGTTGAAACTGCGCGACGCGAC 360  
 QY 2200 CGCACAGATATGCGGGGATCGCCCTTACGCGCAACTTTCCGCGAGCGGACGCCGTA 2259  
 DB 361 CGCACAGATATGCGGGGATCGCCCTTACGCGCAACTTTCCGCGAGCGGACGCCGTA 420  
 QY 2260 CAGCATGCGAATGCGCGGACGCGTGTACGCATCTTCAACAGTCTCGCGCTACGCTCTAT 2319  
 DB 421 CAGCATGCGAATGCGCGGACGCGTGTACGCATCTTCAACAGTCTCGCGCTACGCTCTAT 480  
 QY 2320 GCCGACAGTACCGCGGCCCATCCGATATGCAAGGACCGCGCTGAAAGCCGTATCGGAC 2379  
 DB 481 GCCGACAGTACCGCGGCCCATCCGATATGCAAGGACCGCGCTGAAAGCCGTATCGGAC 540  
 QY 2380 GGGTTGGACCAACAGCTACGGGTCTGCGGTCTATCGGCGAAACCCACAGACAGCGTGA 2439  
 DB 541 GGGTTGGACCAACAGCGGACGCGGTCTGCGGTCTATCGGCGAAACCCACAGACAGCGTGA 600  
 QY 2440 AGTGGGACAGGCGGTGTTGAAGGCAAAATGCGGCGAGTACCCAAACCTCGGCATT 2499  
 DB 601 AGTGGGACAGGCGGTGTTGAAGGCAAAATGCGGCGAGTACCCAAACCTCGGCATT 660  
 QY 2500 GCGCGAAAAACCGGCAAAATACGACAGCGCGCCACACTGGGATGGGACACAGCAC 2559  
 DB 661 GCGCGAAAAACCGGCAAAATACGACAGCGCGCCACACTGGGATGGGACACAGCAC 720  
 QY 2560 TGGAGCGAAAAACAGTGCAAATGCAAAACCCGACAGATTAAGTCTGTTGACAGGCATACGG 2619  
 DB 721 TGGAGCGAAAAACAGTGCAAATGCAAAACCCGACAGATTAAGTCTGTTGACAGGCATACGG 780  
 QY 2620 CACGATGCGGGCGATATCGGCTATCTCAAGGCGCTTCTCTACGAGCGCTACAAAAAC 2679  
 DB 781 CACGATGCGGGCGATATCGGCTATCTCAAGGCGCTTCTCTACGAGCGCTACAAAAAC 840  
 QY 2680 AGCATCAGCGCGACACCGGTGCGGACGCAACATGCGGAAGGAGCGGTCAACGGCAGCTG 2739  
 DB 841 AGCATCAGCGCGACACCGGTGCGGACGCAACATGCGGAAGGAGCGGTCAACGGCAGCTG 900  
 QY 2740 ATGCACTGCGGGCGACTGGGCGGTGTCAACGTTTCGTTGCCGCAACGGGAGATTGAGC 2799  
 DB 901 ATGCACTGCGGGCGACTGGGCGGTGTCAACGTTTCGTTGCCGCAACGGGAGATTGAGC 960  
 QY 2800 GTGGAAGCGGTCTCGCTTACGACCTGTCTCAACAGGATGCAATTCGCCGCAAAAGGCACT 2859  
 DB 961 GTGGAAGCGGTCTCGCTTACGACCTGTCTCAACAGGATGCAATTCGCCGCAAAAGGCACT 1020  
 QY 2860 GCTTTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGTCGGACTCGCGGTCTG 2919  
 DB 1021 GCTTTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGTCGGACTCGCGGTCTG 1080  
 QY 2920 AAGCTGTCCACACCTTGAAGCATTAAGCCGTCCTGTTTGAACGGCGGCGGTGGAACGC 2979  
 DB 1081 AAGCTGTCCACACCTTGAAGCATTAAGCCGTCCTGTTTGAACGGCGGCGGTGGAACGC 1140  
 QY 2980 GACCTGAACGAGCGACACTACAGGTAAACGGGCGCTTTACCGGCGGACTGCAGCAAC 3039  
 DB 1141 GACCTGAACGAGCGACACTACAGGTAAACGGGCGCTTTACCGGCGGACTGCAGCAAC 1200  
 QY 3040 GGCAAGACGGGGCAGCAATATGCGGCACACCCGCTGTTGCCGCTTGGCGCGGAT 3099  
 DB 1201 GGCAAGACGGGGCAGCAATATGCGGCACACCCGCTGTTGCCGCTTGGCGCGGAT 1260  
 QY 3100 GTCGAATTCGGAACAGGCTGGAAACGGCTGGACAGCTTACAGCTACGCCGCTTCCAAACAG 3159  
 DB 1261 GTCGAATTCGGAACAGGCTGGAAACGGCTGGACAGCTTACAGCTACGCCGCTTCCAAACAG 1320  
 QY 3160 TACGCAACACAGCGGACGAGTTCGCGGTAGGCTACCGGTTCTGA 3204  
 DB 1321 TACGCAACACAGCGGACGAGTTCGCGGTAGGCTACCGGTTCTGA 1365



## RESULT 22

AAZ53299  
ID AAZ53299 standard; DNA; 1365 BP.XX  
AC AAZ53299;XX  
DT 21-MAR-2000 (first entry)XX  
DE Neisseria gonorrhoeae ORF 140 partial DNA sequence SEQ ID NO:547.XX  
KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;  
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;  
KW antibacterial; gene therapy; ds.XX  
OS Neisseria gonorrhoeae.XX  
PN W09957280-A2.XX  
PD 11-NOV-1999.XX  
PF 30-APR-1999; 99WO-US09346.XX  
PR 01-MAY-1998; 98US-0083758.XX  
PR 31-JUL-1998; 98US-0094869.XX  
PR 02-SEP-1998; 98US-0098994.XX  
PR 09-OCT-1998; 98US-0099062.XX  
PR 09-OCT-1998; 98US-0103749.XX  
PR 09-OCT-1998; 98US-0103794.XX  
PR 25-FEB-1999; 99US-0121528.XX  
PA (CHIR ) CHIRON CORP.XX  
PA (GENO-) INST GENOMIC RES.XX  
PI Fraser C, Galeotti C, Grandi G, Hickey E, Massignani V, Mora M;  
PI Petersen J, Pizza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;  
PI Tettelin H, Venter JC;XX  
DR WPI; 2000-062150/05.XX  
DR P-PSDB; AAY74537.XX  
PT Novel Neisserial polypeptides predicted to be useful antigens for  
PT vaccines and diagnostics -XX  
PS Claim 7; Page 394; 1453pp; English.XX  
CC AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941  
CC represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides  
CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ5473 represent  
CC PCR primers used in the exemplification of the present invention. The  
CC polypeptides, the polynucleotides, antibodies and compositions of  
CC the invention can be used as vaccines, as diagnostic reagents, and as  
CC immunogenic compositions. The polypeptides can be used in the  
CC manufacture of medicaments for treating or preventing infection due to  
CC Neisserial bacteria (e.g. meningitis and septicaemia), to detect the  
CC presence of Neisseria bacteria, or to raise antibodies. They may also  
CC be used to screen for agonists or antagonists, which may themselves  
CC have use as antibacterial agents. The polynucleotides of the invention  
CC may also be used in gene therapy protocols.XX  
SQ Sequence 1365 BP; 324 A; 405 C; 416 G; 220 T; 0 other;XX  
Query Match 40.2%; Score 1288.2; DB 21; Length 1365;  
Best Local Similarity 96.5%; Pred. No. 0;  
Matches 1317; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 1840 ATGTGGCAGCGGCAAGGGGCGAGGTATCTCAACCGTACCGGACAACGTGTTCCCTTC 1899

DB 1 ATGTGGCAGCGGCAAGGGGCGAGGTATCTCAACAGTACCGGAGACATGTTCCCTTC 60

QY 1900 CTGAGTGGCGGCAAAATCGGGGGGATTTATCTTTCACAAAACATCGAAACGACGGT 1959  
|||||

Db	61	CTGAGTGGCGGCAAAATCGGGCAGGATTATTCTTTCAAAAATATCAAAACCGACGCG	120
QY	1960	GGTCTGCTGGCTTCCCTCGACAGCGTCGAAAAACAGCGGCGAGTGAAGCGCACAGCTG	2019
Db	121		
QY	2020	TCCTATTATGTCGGTCGGGCAATCGGCACGAGCTGCTTTCGGCAGCGGCACATTCCGCG	2079
Db	181		
QY	2080	CCGCGCGGTCTGAACACACGCCGTAGAACAGGGCGGAGCAATCTGGAAAACTGATGGTC	2139
Db	241	CCGCGCGGTCTGAACACACGCCGTAGAACAGGGCGGAGCAATCTGGAAAACTGATGGTC	300
QY	2140	GAACCTGGATGCTCCGGAATCATCCGCAACACCCGAGACGGTTGAACCTCGGCGCCGAC	2199
Db	301		
QY	2200	CGCACAGATATCGCGGGCATCCGCCCTACGGCGCAACTTTCCGGCAGCGGCAGCGGTA	2259
Db	361	CGCACAGATATCGCGGGCATCCGCCCTACGGCGCACAACTTTCCGCACACGCGCGGTA	420
QY	2260	GAGCATGCGAATCGCGCCGACGGGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTAT	2319
Db	421		
QY	2320	GCGCAGTACCGCGCCCATGCCGATATGAGGCGGCGGCGTGAAGCGGCGTATCGGAC	2379
Db	481	GCGCAGTACCGCGCCCATGCCGATATGAGGCGGCGGCGTGAAGCGGCGTATCGGAC	540
QY	2380	GGGTTGGACACACGCTACGGGTCTGCGGCTCATCGCGCAAAACCCACAGGACGGTGA	2439
Db	541	GGGTTGGACACACGCTACGGGTCTGCGGCTCATCGCGCAAAACCCACAGGACGGTGA	600
QY	2440	AGTGGGAACAGGGCGGTGTTGAAGCGAATAATGCGGGCGAGTACCCAAACCGTCGCAATT	2499
Db	601	AGTGGGAACAGGGCGGTGTTGAAGCGAATAATGCGGGCGAGTACCCAAACCGTACCGCAATT	660
QY	2500	GCGCGAAAACCGGCAAAATACGACAGCGCGCACACTGGGCATGGAGCAGACACACA	2559
Db	661	GCGCGAAAACCGGCAAAATACGACAGCGCGCACACTGGGCATGGAGCAGACACACA	720
QY	2560	TGGAGCGAAAACAGTGCAAAATGCAAAAACCGACAGCATTAGTCTGTTTTCAGGCGATACGG	2619
Db	721	TGGAGCGAAAACAGTGCAAAATGCAAAAACCGACAGCATTAGTCTGTTTTCAGGCGATACGG	780
QY	2620	CAGATGCGGGCGATATCGGCTATCTCAAGGCGCTGTTCTCTACGAGCGGTACAAAAAC	2679
Db	781	CAGATGCGGGCGATATCGGCTATCTCAAGGCGCTGTTCTCTACGAGCGGTACAAAAAC	840
QY	2680	AGCATCAGCGCAGCAGCGGTGCGGACGAAACATGCGGAGGCGAGTCAACGCGCAGCTG	2739
Db	841	AGCATCAGCGCAGCAGCGGTGCGGATGATATGCGGAGGCGAGTCAACGCGCAGCTG	900
QY	2740	ATGACGTGGGGCGACTGGGCGGTCTCAACGTTTCGGTTCGCGCAACCGGAGATTGAGG	2799
Db	901	ATGACGTGGGGCGACTGGGCGGTCTCAACGTTTCGGTTCGCGCAACCGGAGATTGAGG	960
QY	2800	GTCGAAGCGGTCTCGGCTACGACCTGCTCAAAACAGGATGCAATTCGCGCAAAAGCAGT	2859
Db	961	CTTGAAGCGGTCTCGGCTACGACCTGCTCAAAACAGGATGCAATTCGCGCAAAAGCAGT	1020
QY	2860	GCTTTGGGCTGGAGCGGCAACAGCCTCACTGAAGCAGACACTGGTCCGACTCGCGGTCTG	2919
Db	1021	GCTTTGGGCTGGAGCGGCAACAGCCTCACTGAAGCAGACACTGGTCCGACTCGCGGTCTG	1080
QY	2920	AAGCTGTCGCAACCCCTTGAAGCATAAAGCGCTCTGTTTGAACGCGGCGCTGCAACGC	2979
Db	1081	AAGCTGTCGCAACCCCTTGAAGCATAAAGCGCTCTGTTTGAACGCGGCGCTGCAACGC	1140
QY	2980	GACCTGAACGAGCGAGCTACACGCTAACGGCGGCTTTACGGCGGCGAGCTGACGAAAC	3039
Db	1141	GACCTGAACGAGCGAGCTACACGCTAACGGCGGCTTTACGGCGGCGAGCTGACGAAAC	1200

```
Qy 3040 GGCAAGACGGGGCAGCGCAATATGCCCGCACACCCGCCCTGGTTGCCGGTCTGGGCGCGGAT 3099
      |||||||
Db 1201 GGCAAGACGGGTGCACGCAATATGCCCGCACACCCCGCGGTTCGGGTCTGGGTTGGAT 1260
      |||||||
Qy 3100 GTCGAATTCGGCAACGGCTGGAACGGCTTGGCAGCTTACAGCTACGCCGGTTCCAAACAG 3159
      |||||||
Db 1261 GTCGAATTCGGCAACGGCTGGAACGGCTTGGCAGCTTACAGCTACGCCGGTTCCAAACAG 1320
      |||||||
Qy 3160 TAGGCAACACAGCGGACGAGTCGGCGTAGGCTACCGGTTCTGA 3204
      |||||||
Db 1321 TAGGCAACACAGCGGACGAGTAATCGCGTAGGCTACCGGTTCTGA 1365
      |||||||

RESULT 23
AAZ53298
ID AAZ53298 standard; DNA; 526 BP.
XX
AC AAZ53298;
XX
DT 21-MAR-2000 (first entry)
XX
DE Neisseria meningitidis ORF 139 partial DNA sequence SEQ ID NO:545.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW antibacterial; gene therapy; ds.
XX
OS Neisseria meningitidis.
XX
PN WO9957280-A2.
XX
PD 11-NOV-1999.
XX
PF 30-APR-1999; 99WO-US09346.
XX
PR 01-MAY-1998; 98US-0083758.
PR 31-JUL-1998; 98US-0094869.
PR 02-SEP-1998; 98US-0098994.
PR 02-SEP-1998; 98US-0099062.
PR 09-OCT-1998; 98US-0103749.
PR 09-OCT-1998; 98US-0103794.
PR 09-OCT-1998; 98US-0103796.
PR 25-FEB-1999; 99US-0121528.
XX
PA (CHIR ) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX
PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
PI Petersen J, Pizza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
PI Tettelin H, Venter JC;
XX
DR WPI; 2000-062150/05.
DR P-PSDB; AAY74536.
XX
PT Novel Neisserial polypeptides predicted to be useful antigens for
PT vaccines and diagnostics
XX
PS Claim 7; Page 393; 1453pp; English.
XX
CC AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941
CC represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides
CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ5473 represent
CC PCR primers used in the exemplification of the present invention. The
CC polypeptides, the polynucleotides, antibodies and compositions of
CC the invention can be used as vaccines, as diagnostic reagents, and as
CC immunogenic compositions. The polypeptides can be used in the
CC manufacture of medicaments for treating or preventing infection due to
CC Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
CC presence of Neisseria bacteria, or to raise antibodies. They may also
CC be used to screen for agonists or antagonists, which may themselves
CC have use as antibacterial agents. The polynucleotides of the invention
CC may also be used in gene therapy protocols.
```

```
XX
SQ Sequence 526 BP; 159 A; 142 C; 131 G; 94 T; 0 other;
      Query Match 15.4%; Score 494.2; DB 21; Length 526;
      Best Local Similarity 98.7%; Pred. No. 3.1e-127;
      Matches 520; Conservative 0; Mismatches 3; Indels 4; Gaps 2;

Qy 1 ATGCGAAGACGACCCCAACTTCCCTACAAAAAATTTCAAAACCGGTCCATGCGGTTAGCT 60
      |||||||
Db 1 ATGCGAAGACGACCCCAACTTCCCTACAAAAAATTTCAAAACCGGTCCATGCGGTTAGCT 60
      |||||||
Qy 61 GTTGCACACACACTTTCTGCTCTAGGCGGCGGC---GGCGGCACTTCTGCGCGCGAC 117
      |||||||
Db 61 GTTGCACACACACTTTCTGCTCTAGGCGGCGGCGGCGGCACTTCTGCGCGCGAC 120
      |||||||
Qy 118 TTCAATGCAGGCGGCGGCTATCGGCAGCAACAGCAGCAACCAACGCAAAATCAGCA 177
      |||||||
Db 121 TTCAATGCAGGCGGCGGCTATCGGCAGCAACAGCAGGCGCAACAGCAAAATCAGCA 180
      |||||||
Qy 178 GCAGTATCTTACGCGCGGTATCAAGAACGAAATGTGCAAAAGACAGAAGCATGCTCTGTGCC 237
      |||||||
Db 181 GCAATATCTTACGCGCGGTATCAAGAACGAAATGTGCAAAAGACAGAAGCATGCTCTGTGCC 240
      |||||||
Qy 238 GGTCCGGATGACGTTGCGGTTTACAGACAGGATGCCAAATCAATGCCCCCCCCCGAAT 297
      |||||||
Db 241 GGTCCGGATGACGTTGCGGTTTACAGACAGGATGCCAAATCAATG-CGCCGCCCGCAAT 299
      |||||||
Qy 298 CTGCATACCGGAGACTTTACAAACCCCAATGACGCATCAAGAATTTGATCAACCTCAA 357
      |||||||
Db 300 CTGCATACCGGAGACTTTACAAACCCCAATGACGCATCAAGAATTTGATCAACCTCAA 359
      |||||||
Qy 358 CCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTATGTCGATACAGGC 417
      |||||||
Db 360 CCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTATGTCGATACAGGC 419
      |||||||
Qy 418 GAATCCGTCGCGCAGCATATCCTTTCCCGAATGTATGCGCAAAAAGACACGCGCTATAAC 477
      |||||||
Db 420 GAATCCGTCGCGCAGCATATCCTTTCCCGAATGTATGCGCAAAAAGACACGCGCTATAAC 479
      |||||||
Qy 478 GAAATTTACAAAAAATACTACGCGGTATATGCGGAAGGAGCGCCTGA 524
      |||||||
Db 480 GAAATTTACAAAAAATACTACGCGGTATATGCGGAAGGAGCGCCTGA 526
      |||||||

RESULT 24
AAZ53297
ID AAZ53297 standard; DNA; 528 BP.
XX
AC AAZ53297;
XX
DT 21-MAR-2000 (first entry)
XX
DE Neisseria meningitidis ORF 138 partial DNA sequence SEQ ID NO:543.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW antibacterial; gene therapy; ds.
XX
OS Neisseria meningitidis.
XX
PN WO9957280-A2.
XX
PD 11-NOV-1999.
XX
PF 30-APR-1999; 99WO-US09346.
XX
PR 01-MAY-1998; 98US-0083758.
PR 31-JUL-1998; 98US-0094869.
PR 02-SEP-1998; 98US-0098994.
PR 02-SEP-1998; 98US-0099062.
PR 09-OCT-1998; 98US-0103749.
PR 09-OCT-1998; 98US-0103794.
PR 09-OCT-1998; 98US-0103796.
```

PR 25-FEB-1999; 99US-0121528.  
XX (CHIR ) CHIRON CORP.  
PA (GENO-) INST GENOMIC RES.  
PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;  
PI Petersen J, Pizsa M, Rappuoli R, Ratti G, Scalato E, Scarselli M;  
PI Tettelin H, Venter JC;  
XX WPI: 2000-062150/05.  
DR P-PSDB; AAY74535.  
XX Novel Neisserial polypeptides predicted to be useful antigens for  
PT vaccines and diagnostics  
XX Claim 7; Page 393; 1453pp; English.  
XX AA253015 to AA254536, AA254577 to AA254615, and AAY74253 to AAY75941  
CC represent novel Neisseria meningitis and N. gonorrhoea polynucleotides  
CC and polypeptides. AA254537 to AA254576 and AA254616 to AA255473 represent  
CC PCR primers used in the exemplification of the present invention. The  
CC polypeptides, the polynucleotides, antibodies and compositions of  
CC the invention can be used as vaccines, as diagnostic reagents, and as  
CC immunogenic compositions. The polypeptides can be used in the  
CC manufacture of medicaments for treating or preventing infection due to  
CC Neisserial bacteria (e.g. meningitis and septicemia), to detect the  
CC presence of Neisseria bacteria, or to raise antibodies. They may also  
CC be used to screen for agonists or antagonists, which may themselves  
CC have use as antibacterial agents. The polynucleotides of the invention  
CC may also be used in gene therapy protocols.  
XX Sequence 528 BP; 159 A; 142 C; 131 G; 96 T; 0 other;  
SQ

Query Match 14.7%; Score 470.6; DB 21; Length 528;  
Best Local Similarity 98.1%; Pred. No. 1.2e-120;  
Matches 519; Conservative 0; Mismatches 4; Indels 6; Gaps 4;  
QY 1 ATGCGAACACCACTTCTCCCTACAAAACCTTTCAAAACCGCTGCCATGGGTTAGCT 60  
Db 1 ATGCGAACACCACTTCTCCCTACAAAACCTTTCAAAACCGCTGCCATGGGTTAGCT 60  
QY 61 GTTGAACAACACTTCTTCTGCTTACAGAGGATGCGGCGGC---GGCGGCACCTTCTGCGCCGAC 117  
Db 61 GTTGAACAACACTTCTTCTGCTTACAGAGGATGCGGCGGCAGGCGGACCTTCTGCGCCGAC 120  
QY 118 TTCAATGACGGCGGACCGGTATCGGCAGCAGCAGCAGCAGCAGCAATCAGCA 177  
Db 121 TTCAATGACGGCGGATACCGGTATCGGCAGCAGCAGCAGCAGCAATCAGCA 180  
QY 178 GCAGTATCTTACGCCGTATCAAGAACGAAATGTGCAAGACAGAGCATGCTCTGTGCC 237  
Db 181 GCAGTATCTTACGCCGTATCAAGAACGAAATGTGCAAGACAGAGCATGCTCTGTGCC 240  
QY 238 GGTGCGGATGACGTTGCGGTATACAGACAGGATGCGCAATCAATGCCCCCCCCCGAAT 297  
Db 241 GGTGCGGATGACGTTGCGGTATACAGACAGGATGCGCAATCAATGCCCCCCCCCGAAT 299  
QY 298 CTGCATACCGGACGTTTCAACCCAAATGACGCA-TACAAGAATTTGATCAACCTCAA 356  
Db 300 CTGCATACCGGACGTTTCAACCCAAATGACGCAATTAAGAATTTGATCAACCTCAA 359  
QY 357 ACCTGCAATTGAAGCAGGCTATACAGACGCGGGGTAGGTAGGTATCGTCATACAGG 416  
Db 360 ACCTGCAATTGAAGCAGGCTATACAGACGCGGGGTAGGTAGGTATCGTCACACAGG 419  
QY 417 CGAATCCGTCGCGAGCATATCCTTTCCGAACTGTATGCGAGAAAGAACACGGCTATAA 476  
Db 420 CGAATCCGTCGCGAGCATATCCTTTCCGAACTGTATGCGAGAAAGAACACGGCTATAA 479  
QY 477 CGAATATATCGGCGGTATATGCGGAAGGAAGCGGCTGA 524  
Db 480 CGAATATACGAAATAATATACGGCGGTATATGCGGAAGGAAGCGGCTGA 528

RESULT 25  
AA253296  
ID AA253296 standard; DNA; 537 BP.  
XX  
AC AA253296;  
XX  
DT 21-MAR-2000 (first entry)  
XX  
DE Neisseria gonorrhoea ORF 138 partial DNA sequence SEQ ID NO:541.  
XX  
KW Neisseria meningitidis; Neisseria gonorrhoea; antigen; vaccine;  
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;  
KW antibacterial; gene therapy; ds.  
XX  
OS Neisseria gonorrhoea.  
XX  
PN WO9957280-A2.  
XX  
PD 11-NOV-1999.  
XX  
PF 30-APR-1999; 99WO-US09346.  
XX  
PR 01-MAY-1998; 98US-0083758.  
PR 31-JUL-1998; 98US-0094869.  
PR 02-SEP-1998; 98US-0098994.  
PR 02-SEP-1998; 98US-0099062.  
PR 09-OCT-1998; 98US-0103749.  
PR 09-OCT-1998; 98US-0103794.  
PR 09-OCT-1998; 98US-0103796.  
PR 25-FEB-1999; 99US-0121528.  
XX  
PA (CHIR ) CHIRON CORP.  
PA (GENO-) INST GENOMIC RES.  
XX  
PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;  
PI Petersen J, Pizsa M, Rappuoli R, Ratti G, Scalato E, Scarselli M;  
PI Tettelin H, Venter JC;  
DR WPI: 2000-062150/05.  
DR P-PSDB; AAY74534.  
XX  
PT Novel Neisserial polypeptides predicted to be useful antigens for  
PT vaccines and diagnostics  
XX  
PS Claim 7; Page 392; 1453pp; English.  
XX  
CC AA253015 to AA254536, AA254577 to AA254615, and AAY74253 to AAY75941  
CC represent novel Neisseria meningitis and N. gonorrhoea polynucleotides  
CC and polypeptides. AA254537 to AA254576 and AA254616 to AA255473 represent  
CC PCR primers used in the exemplification of the present invention. The  
CC polypeptides, the polynucleotides, antibodies and compositions of  
CC the invention can be used as vaccines, as diagnostic reagents, and as  
CC immunogenic compositions. The polypeptides can be used in the  
CC manufacture of medicaments for treating or preventing infection due to  
CC Neisserial bacteria (e.g. meningitis and septicemia), to detect the  
CC presence of Neisseria bacteria, or to raise antibodies. They may also  
CC be used to screen for agonists or antagonists, which may themselves  
CC have use as antibacterial agents. The polynucleotides of the invention  
CC may also be used in gene therapy protocols.  
XX  
SQ Sequence 537 BP; 169 A; 136 C; 130 G; 102 T; 0 other;  
Query Match 13.5%; Score 433.4; DB 21; Length 537;  
Best Local Similarity 93.2%; Pred. No. 2.8e-110;  
Matches 504; Conservative 0; Mismatches 16; Indels 21; Gaps 4;  
QY 1 ATCGCAACGACCCCAACCTTCCCTACAAAACCTTTCAAAACCGCTGCCATGGGTTAGCT 60  
Db 1 ATCGCAACGACCCCTCAACCTTCCCTACAAAACCTTTCAAAACCGCTGCCATGGGTTAGCT 60  
QY 61 GTTGAACAACACTTCTTCTGCTCTTAGGCGCGGC---GGCGGCACCTTCTGCGCCGAC 117  
|||||

Db 61 GTTGCAACAACACTTTCTGCTGCTTAGGGGGGGGGAGCGGCGACTTCTGTCCCGAC 120  
 Qy 118 TTCAATGACGGCGCACCGGTATCGGCAGCAACACAGACAGCAACAAACGCGAAATCAGCA 177  
 Db 121 TTTAATGACGGCGCACCGGTATCGGCAGCAACACAGACAGCAACGATAGCGGATCAGCA 180  
 Qy 178 GCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAGACAGCAAGCATGCTCTGTGCC 237  
 Db 181 GCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAGACAGCAAGCATGCTCTGTGCC 240  
 Qy 238 GGTGGGATGACGTGGGTATACAGACAGGGATGCCAAATCAATGCCCGCCCGCAAT 297  
 Db 241 GGTGGGATGACGTGGGTATACAGACAGGGATGCCAAATCAATGCCCGCCCGCAAT 296  
 Qy 298 CTGATACCGGAGACTTTTACAAACCCAAATGACGCATATCAAGAAATTTGATCAACCTCAA 356  
 Db 297 CTGATACCGGAGACTTTTCAAAACCCAAATGACGCATATCAAGAAATTTGATCAACCTCAA 356  
 Qy 357 ACCTGCAATTGAAGCAGCTATACAGGACGCGGGGTAGAGTAGGTATCGTCGATACAGG 416  
 Db 357 ACCTGCAATTGAAGCAGCTATACAGGACGCGGGGTAGAGTAGGTATCGTCGATACAGG 416  
 Qy 417 CGAATCCGTCGGCAGCATATCTTTCCGGAATGTATGGCAGAGAAAGACGCGCTATAA 476  
 Db 417 CGAATCCGTCGGCAGCATATCTTTCCGGAATGTATGGCAGAGAAAGACGCGCTATAA 476  
 Qy 477 CGA-----AAATTACAAAACACTATACGGCGTATATGCGGAAGGAAGCGCGTG 523  
 Db 477 CGAATAATACAAAACAAATTTACAAAACACTATACGGCGTATATGCGGAAGGAAGCGCGTG 536  
 Qy 524 A 524  
 Db 537 A 537

RESULT 26  
 AAV03603  
 ID AAV03603 standard; DNA; 286 BP.  
 XX  
 AC AAV03603;  
 DT 22-OCT-1998 (first entry)  
 XX  
 DE Neisseria meningitidis DNA sequence E85.  
 KW N. gonorrhoeae; N. lactamica; chromosome Z2491; region 1; region 2;  
 KW region 3; pathogenicity; blood-brain barrier; diagnosis; infection;  
 KW meningitis; ss.  
 XX  
 OS Neisseria meningitidis.  
 PN W09802547-A2.  
 XX  
 PD 22-JAN-1998.  
 XX  
 PF 11-JUL-1997; 97WO-FR01295.  
 XX  
 PR 12-JUL-1996; 96FR-0008768.  
 XX  
 PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.  
 PA (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.  
 PA (SMIK ) SMITHKLINE BEECHAM.  
 XX  
 PI Nassif X, Tinsley C, Achtman M, Merker P, Ruelle J;  
 PI Vinals C;  
 XX  
 DR WPI; 1998-110594/10.

Genes present in Neisseria meningitidis but not other Neisseria  
 PT species - and related host cells, RNA, anti-sense sequences,  
 PT polypeptide(s) and antibodies, useful for diagnosing Neisseria  
 PT meningitidis infection and in protective vaccines  
 XX

PS Example 4; Page 128; 150pp; French.  
 XX  
 CC AAV03575-606 represent sequences that are present in Neisseria  
 CC meningitidis and N. gonorrhoeae but not in N. lactamica, except for the  
 CC genes involved in biosynthesis of the capsule polysaccharide, frrA or C,  
 CC opc, porA, rotamase, sequence Icl106, Iga protease, pillin, pilC,  
 CC proteins which bind transferrin and opacity proteins. The DNA sequences  
 CC are responsible for the differences in pathogenicity between N.  
 CC meningitidis and N. gonorrhoeae, specifically they include the genes that  
 CC allow N. meningitidis to cross the blood-brain barrier. DNA sequences  
 CC common to N. meningitidis and N. gonorrhoeae, but absent from N.  
 CC lactamica, are responsible for colonisation and penetration of the  
 CC mucosa. The DNA sequences can be used to produce probes and primers, and  
 CC antibodies produced against the encoded proteins are used in standard  
 CC hybridisation/immunoassay processes for diagnosis of N. meningitidis  
 CC infection, particularly meningitis.  
 XX  
 SQ Sequence 286 BP; 88 A; 73 C; 63 G; 62 T; 0 other;

Query Match 8.5%; Score 270.8; DB 19; Length 286;  
 Best Local Similarity 99.0%; Pred. No. 4e-65;  
 Matches 283; Conservative 0; Mismatches 2; Indels 1; Gaps 1;  
 Qy 865 AATTCGGAGGAGCAGTACCGCCAAAGCGTTGCTCGCCTATTCCGGCGGTGATAAACACAGAC 924  
 Db 1 AATTCGGAGGAGCAGTACCGCCAAAGCGTTGCTCGCCTATTCCGGCGGTGATAAACACAGAC 60  
 Qy 925 GAGGTTATCCGCCGTGATGCAACAGAGCGATTACGCGCAACTTGTCTACACATCCGTAAT 984  
 Db 61 GAGGTTATCCGCCGTGATGCAACAGAGCGATTACGCGCAACTTGTCTACACATCCGTAAT 120  
 Qy 985 AAAACATGCTTTTTCATTTTTCGGCAAGCAATGACGCAAGCTCAGCCCAACAC-ACT 1043  
 Db 121 AAAACATGCTTTTTCATTTTTCGGCAAGCAATGACGCAAGCTCAGCCCAACACAACT 180  
 Qy 1044 GACCTATTGCCATTTTATGAAAAGATGCTCAAAAAGGCATTATCAGATCGCAGCGGT 1103  
 Db 181 GACCTATTGCCATTTTATGAAAAGATGCTCAAAAAGGCATTATCAGATCGCAGCGGT 240  
 Qy 1104 AGACGCGAGTGGAGAAAAGTTCAATGGCTCCCAACCATTTGCCGAATT 1149  
 Db 241 AGACGCGAGTGGAGAAAAGTTCAATGGCTCCCAACCATTTGCCGAATT 286

RESULT 27  
 AAA15396  
 ID AAA15396 standard; DNA; 286 BP.  
 XX  
 AC AAA15396;  
 DT 04-SEP-2000 (first entry)  
 XX  
 DE Genomic fragment of Neisseria meningitidis Z2491.  
 XX  
 KW Pathogenic strain; Neisseria; vaccine; Neisseria infection; ss.  
 OS Neisseria meningitidis.  
 PN W0200026375-A2.  
 XX  
 PD 11-MAY-2000.  
 XX  
 PF 28-OCT-1999; 99WO-FR02643.  
 XX  
 PR 30-OCT-1998; 98FR-0013693.  
 XX  
 PA (INRM ) PASTEUR MERIEUX SERUMS & VACCINS SA.  
 PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.  
 XX  
 PI Aujame L, Bouchardon A, Renauld-Mongenie G, Rokbi B, Nassif X;  
 PI Tinsley C, Perrin A;  
 XX  
 DR WPI; 2000-365622/31.

XX New polypeptide specific for pathogenic *Neisseria* useful in therapeutic  
PT or preventative vaccines and for diagnosis -  
XX  
XX  
XX Claim 1; Page 44; 187pp; French.

XX The present sequence represents a genomic fragment of *Neisseria*  
CC meningitidis. The specific amino acid sequence of the protein that is specific  
CC for pathogenic strains of *Neisseria*. The polynucleotides, polypeptides,  
CC or their antigenic fragments, are used in vaccines to treat or protect  
CC against *Neisseria* infections, particularly by *N. meningitidis*. The  
CC polynucleotide sequences are also used for recombinant production of  
CC the polypeptide and to produce attenuated *Neisseria* strains that  
CC overexpress it, or express it in a non-toxic mutant form.  
CC note: the present sequence is included in Claim 1, but it is those  
CC sequences that do not include the present sequence that are actually  
CC claimed.

XX Sequence 286 BP; 88 A; 73 C; 63 G; 62 T; 0 other;

Query Match 8.5%; Score 270.8; DB 21; Length 286;  
Best Local Similarity 99.0%; Pred. No. 4e-65;  
Matches 283; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 865 AATTCGAGGAGCAGTACCCCAAGCGTTGCTCGCTATTCCGGCGGTGATAAAACAGAC 924

DB 1 AATTCGAGGAGCAGTACCCCAAGCGTTGCTCGCTATTCCGGCGGTGATAAAACAGAC 60

QY 925 GAGGGTATCCGCTGATGCAACAGAGGATACGGCAACTTGCTTACCACATCCCGTAAT 984

DB 61 GAGGGTATCCGCTGATGCAACAGAGGATACGGCAACTTGCTTACCACATCCCGTAAT 120

QY 985 AAAAATGCTTTTTCATTTTTCGGCAAGCAATGACGCAAGCTCAGCCCAACAC-ACT 1043

DB 121 AAAAATGCTTTTTCATTTTTCGGCAAGCAATGACGCAAGCTCAGCCCAACACAACT 180

QY 1044 GACCTATTGCCATTTTATGAAAAGATGCTCAAAAGGCAATTATCACATCGCAGCGT 1103

DB 181 GACCTATTGCCATTTTATGAAAAGAGCGCTCAAAAGGCAATTATCACATCGCAGCGT 240

QY 1104 AGACCCGATGGAGAAAGTTCAATGCTCCACCACTTGGCGAATT 1149

DB 241 AGACCCGATGGAGAAAGTTCAATGCTCCACCACTTGGCGAATT 286

RESULT 28

AAA82003

ID AAA82003 standard; DNA; 577 BP.

XX AAA82003;

XX 04-DEC-2000 (first entry)

DE N. meningitidis partial DNA sequence gnm\_550 SEQ ID NO:550.

XX *Neisseria meningitidis*; *Neisseria gonorrhoeae*; genome; immunogenic;  
KW antigen; vaccine; diagnosis; infection; antibacterial; identification;  
KW Meningococcus B; MenB; ds.  
XX *Neisseria meningitidis*.

XX WO200022430-A2.

XX 20-APR-2000.

XX 08-OCT-1999; 99WO-US23573.

XX 09-OCT-1998; 98US-0103794.

XX 30-APR-1999; 99US-0132068.

XX (CHIR ) CHIRON CORP.

XX Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;

PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scariato V;  
PI Rappuoli R, Pizza M;  
XX WPI; 2000-318079/27.

XX Isolated nucleotide sequences of *Neisseria meningitidis* which can be  
PT used in the diagnosis and treatment of *N. meningitidis* infection and  
PT other *Neisseria* infections, for example, *N. gonorrhoea* -  
XX  
XX Claim 7; Page 1652-1653; 1760pp; English.

XX The present invention describes methods of obtaining immunogenic  
CC proteins from *Neisseria* genomic sequences. AAA81453 to AAA82414  
CC represent specifically claimed *Neisseria meningitidis* genomic DNA  
CC sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent  
CC *Neisseria* DNA sequences and their corresponding proteins; AAA81254 to  
CC AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the  
CC isolation of *Neisseria meningitidis* DNA sequences; and AAA81322 to  
CC AAA81452 represent *Neisseria meningitidis* MenB polynucleotide ORF  
CC sequences, which are all used in the exemplification of the present  
CC invention. The nucleic acid sequences, protein sequences, and antibodies  
CC against them, can be used in the manufacture of a composition. The  
CC composition can be used as a medicament (or in the manufacture of a  
CC medicament) for treating, preventing or diagnosing infection due to  
CC *Neisseria* bacteria. For example, some of the identified proteins could  
CC be components of vaccines against Meningococcus B; against all serotypes;  
CC and/or against all pathogenic *Neisseria*. Identification of sequences  
CC from the bacterium will also facilitate production of biological probes,  
CC particularly organism-specific probes. Attempts to make efficacious  
CC Meningococcus B vaccines have failed mainly due to antigen tolerance.  
CC Multivalent vaccines have also been tried but none have successfully  
CC overcome antigenic variability. The provision of further, complete  
CC sequences may provide an opportunity to identify secreted or surface  
CC exposed proteins that may be presumed targets for the immune system and  
CC which are not antigenically variable or at least more conserved than  
CC other more variable regions.

XX Sequence 577 BP; 189 A; 134 C; 99 G; 155 T; 0 other;

Query Match 2.3%; Score 74.4; DB 21; Length 577;  
Best Local Similarity 92.9%; Pred. No. 3e-10;  
Matches 78; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 ATCGCAACGACCCCAACCTTCCCTACAAAACCTTTCAAAACGGCTGCCATGGCTTAGCT 60

DB 494 ATGTCAACGACTCCACCTTCCCTACACACACTTTCAAAACGACTGCCATGGCTTAGCT 553

QY 61 GTTGCAACAACACTTTCTGCTGC 84

DB 554 GTTGCAACAACACTTTCTGCTGC 577

RESULT 29

AAA81823/c

ID AAA81823 standard; DNA; 476 BP.

XX AAA81823;

XX 04-DEC-2000 (first entry)

DE N. meningitidis partial DNA sequence gnm\_370 SEQ ID NO:370.

XX *Neisseria meningitidis*; *Neisseria gonorrhoeae*; genome; immunogenic;  
KW antigen; vaccine; diagnosis; infection; antibacterial; identification;  
KW Meningococcus B; MenB; ds.  
XX *Neisseria meningitidis*.

XX WO200022430-A2.

XX 20-APR-2000.

XX 08-OCT-1999; 99WO-US23573.

```
XX 09-OCT-1998; 98US-0103794.
PR 30-APR-1999; 99US-0132068.
XX (CHIR ) CHIRON CORP.
XX Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;
PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
PI Rappuoli R, Pizza M;
XX WPI: 2000-318079/27.
XX Isolated nucleotide sequences of Neisseria meningitidis which can be
PT used in the diagnosis and treatment of N. meningitidis infection and
PT other Neisserial infections, for example, N.gonorrhoea -
XX Claim 7; Page 1608; 1760pp: English.
XX The present invention describes methods of obtaining immunogenic
CC proteins from Neisseria genomic sequences. AAA81453 to AAA82414
CC represent specifically claimed Neisseria meningitidis genomic DNA
CC sequences: AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
CC Neisseria DNA sequences and their corresponding proteins; AAA81254 to
CC AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the
CC isolation of Neisseria meningitidis DNA sequences; and AAA81322 to
CC AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF
CC sequences, which are all used in the exemplification of the present
CC invention. The nucleic acid sequences, protein sequences, and antibodies
CC against them, can be used in the manufacture of a composition. The
CC composition can be used as a medicament (or in the manufacture of a
CC medicament) for treating, preventing or diagnosing infection due to
CC Neisserial bacteria. For example, some of the identified proteins could
CC be components of vaccines against Meningococcus B; against all serotypes;
CC and/or against all pathogenic Neisseriae. Identification of sequences
CC from the bacterium will also facilitate production of biological probes,
CC particularly organism-specific probes. Attempts to make efficacious
CC Meningococcus B vaccines have failed mainly due to antigen tolerance.
CC Multivalent vaccines have also been tried but none have successfully
CC overcome antigenic variability. The provision of further, complete
CC sequences may provide an opportunity to identify secreted or surface
CC exposed proteins that may be presumed targets for the immune system and
CC which are not antigenically variable or at least more conserved than
CC other more variable regions.
XX Sequence 476 BP; 84 A; 154 C; 125 G; 111 T; 2 other;
SQ Query Match 2.0%; Score 65.4; DB 21; Length 476;
Best Local Similarity 70.2%; Pred. No. 8.e-08;
Matches 87; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
QY 1063 GAAAAGATGCTCAAAAGGCGATTATCAGTCGCGAGGCGTAGACCGGAGGAAAG 1122
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
476 GAAAAGAGCGCTCAAAAGGCGATTATCAGTCGCGAGGCGTAGACCGGAGGAAAG 417
QY 1123 TTCAATGGCTCCAAACATTTGCGGATTAATCTGCGCATGTGTCCTATCGGCACCCCTATGAA 1182
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
416 TTCAATATCCGTTTGACGCGGTGCGGAGCTGACGCGCTGTCGCGACTCGCACCCNCTGTC 357
QY 1183 GCAA 1186
DB ||||
356 GGAA 353
RESULT 30
ID AAA10594
XX AAA10594 standard; DNA; 10732 BP.
AC AAA10594;
XX 29-JUN-2000 (first entry)
DT YTHRYSGSRTHRRGGYSGRVARASHSSRGNTHTHRGTHRRGVAASMTGRMTS 9745
XX Gene encoding a subunit of cellulose synthase.
DE GCAACCCCTTGAGCGATAAAGCCGCTCTCTTTGCAACGCGCGGCGTGGACACGCGACCTGAA 2987
XX
```

```
KW Cellulose synthase; cellulose production; increase yield; ds.
XX Vigna angularis.
OS JP2000060568-A.
PN 29-FEB-2000.
PD 26-AUG-1998; 98JP-0239998.
XX 26-AUG-1998; 98JP-0239998.
XX (MIZU/) MIZUNO K.
PA (OUIP ) OJI PAPER CO.
XX WPI: 2000-342371/30.
DR P-PSDB; AAY85179.
XX A gene encoding a cellulose synthetic equipment - for the improvement
PT in the amount of cellulose synthesised in a plant body
XX Claim 2; Page 14-21; 32pp; Japanese.
XX This sequence represents a gene encoding a subunit of the cellulose
CC synthase complex of Vigna angularis. The invention relates to subunits of
CC cellulose synthetic equipment, that can be used to increase the amount of
CC cellulose synthesised by a plant. The proteins and genes encoding them
CC can also be used to improve the properties of the cellulose being
CC produced by a plant.
XX Sequence 10732 BP; 3149 A; 1212 C; 2074 G; 2046 T; 2251 other;
SQ Query Match 1.9%; Score 62; DB 21; Length 10732;
Best Local Similarity 13.9%; Pred. No. 3.8e-06;
Matches 98; Conservative 300; Mismatches 304; Indels 3; Gaps 2;
QY 2388 CCACAACCTACGGGTCTCGCGTCATCGCCAAACCAACAGGAGCGTGGAACTGGGA 2447
DB :|||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
9208 HSGYAAMTSRAGGSRMTGYSGASASNTHRSTHRTYRGVAAGVAGYSGGARGASNGGAR 9267
QY 2448 ACAGGCGCGGTGTGAAGCAAAATCGCGCAGTAGTACCCAAACCCGTCGCGCATTCGCCGCAA 2507
DB :|||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
9268 GHSRARGARGTHRAAGVASRHSYSGNGSNTHRYSYSAAYSGGCGYSGNARGARGVAAR 9327
QY 2508 AACCGCGGAAATACGACAGCAGCGCCACACTGGCGATGGACACACGACGACGACGCA 2567
DB :|||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
9328 GYSARGRGYRAAAAYASNTYSASNGVAGMTGYARGASRHRGARGARGSRYSSTHRSN 9387
QY 2568 AAACAGTCAATGCAAAACCGACAGCATTAGTCTGTTGCGAGGCATACGCGCAGTCGC 2627
DB :|||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
9388 SRVAGSRVAASTHRSRRGTHRRTHRYASNTHTHRCGNCYSAAVAGGASNSYVAYS 9447
QY 2628 GGGCGATATCGGCTATCTCAAAGGCGCTGTTCTCCACGAGCCTACAAAACACGATCAG 2687
DB :|||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
9448 GSRASNARYSASNGNSHRARVAMTHRARGTHRAASRYSHSGSRGNGASHSYGHS 9507
QY 2688 CCGCAGACCGGTCGGGACGACATCGGGAAGCAGCGCTCAACGCGACGCTGATCGAGCT 2747
DB :|||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
9508 SRVAGSHSSRASNAASRHSRHSRAASRAAAGYSRASASYSYTHSRCYSAAGSR 9567
QY 2748 GGGCGCAGTCGGCGGTGTCAACGTTCCGTCGCAACGGAGGATTTTCACGCTCGAAGG 2807
DB :|||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
9568 TRAAASRAASRGA - SHSHARGSRGYSYSGRSRCYSYSSRVAGYAAASRASASNTASAS 9625
QY 2808 CGGTCTGCGCTACGACCTCTCAAACAGGATGCTTCGCCGCAAAAGGAGGAGTCTTTGGG 2867
DB :|||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
9626 HVAGMTGYSAAVAVASRVAGYSAAATHRGNSRRSRAASRYSAAVATHRGASNSRSGTRHG 9685
QY 2868 CTGGAGCGGCAACAGCCTCACTGAGGACACTGGTCGCGACTCGCGGGTCTGAAGCTGTC 2927
DB :|||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
9686 YTHRYSGSRTHRRGGYSGRVARASHSSRGNTHTHRGTHRRGVAASMTGRMTS 9745
QY 2928 GCAACCCCTTGAGCGATAAAGCCGCTCTCTTTGCAACGCGCGGCGTGGACACGCGACCTGAA 2987
XX
```

Db 9746 RASHSSRASSRYSRYSSANTHRCYSGSHYSGYTHRVARGY-TRGNASVAVAYSTHRG 9804  
QY 2988 CGGACCGGACTACACGCTAACGGCGGCTTACCGCGCGGACTGCAGCAACCGGCAAGAC 3047  
Db 9805 GNASHSVATHRHSYSGYRASASHSASYVAAAHGHSVAASHASHASHSRAR 9864  
QY 3048 GGGGCGACCAATATGCCGACACCGCGCTGGTTGCGGTCTGGG 3092  
Db 9865 GGYSRVAHSASTHRGNASNRNGNCYSHSCYSSRCSRRSRV 9909  
RESULT 31  
AAT91476  
ID AAT91476 standard; DNA; 913 BP.  
XX AC AAT91476;  
XX DT 07-JAN-1998 (first entry)  
XX DE Mycobacterium tuberculosis antigen Tbra32 encoding DNA.  
XX KW Antigen; immunogen; vaccine; tuberculosis; non specific adjuvant;  
XX KW skin testing; M.tuberculosis; ss.  
XX OS Mycobacterium tuberculosis.  
XX FH Key Location/Qualifiers  
XX CDS 3..208  
XX FT /\*tag= a  
XX FT /product= Antigen\_Tbra32  
XX FT /transl\_except= (pos:126..127,aa:Pro)  
XX FT /note= "Pro coded for by cc; no stop codon given"  
XX PN WO9709428-A2.  
XX PD 13-MAR-1997.  
XX PF 30-AUG-1996; 96WO-US14674.  
XX PR 12-JUL-1996; 96US-0680574.  
XX PR 01-SEP-1995; 95US-0523436.  
XX PR 22-SEP-1995; 95US-0533634.  
XX PR 22-MAR-1996; 96US-0620874.  
XX PR 05-JUN-1996; 96US-0659683.  
XX PA (CORI-) CORIXA CORP.  
XX PI Campos-neto A, Dillon DC, Houghton R, Reed SG, Skeiky YAW;  
XX PI Twardzik DR, Vedwick TH;  
XX DR WPI: 1997-192903/17.  
XX DR P-PSDB; AAW32434.  
XX PT New immunogenic polypeptide(s) from Mycobacterium tuberculosis - are  
XX PT useful in vaccines for prevention or treatment of tuberculosis, also  
XX PT for diagnosis  
XX PS Claim 3; Page 62-63; 168pp; English.  
XX CC A new immunogenic polypeptide has been developed comprising an  
XX CC immunogenic part of a soluble Mycobacterium tuberculosis antigen (or  
XX CC its variant differing only in conservative substitutions and/or  
XX CC modifications). The present sequence encodes a specifically claimed  
XX CC M.tuberculosis antigen, Tbra32. The immunogenic protein, and fusion  
XX CC proteins containing one or more of the proteins or one of the proteins  
XX CC plus ESAT-6, are useful in vaccines, preferably when formulated with a  
XX CC non-specific adjuvant, to induce an immune response against  
XX CC M.tuberculosis (for treatment or prevention).  
SQ Sequence 913 BP; 95 A; 417 C; 278 G; 118 T; 5 other;  
Query Match 1.5%; Score 49.6; DB 18; Length 913;

Best Local Similarity 47.7%; Pred. No. 0.0031;  
Matches 136; Conservative 1; Mismatches 148; Indels 0; Gaps 0;  
QY 2065 GCGGCGACATCCCGCGCGCGCGGTCTGAACACAGCGCGTAGAACAGGCGGCGCAATCTG 2124  
Db 605 GCGCGCGAGNGCGTGCCCGCGCGCGCGCGCGCAAGCGCCAAAGCCGCGGTTCGCCACCGG 664  
QY 2125 GAAACCTGATGCTCGAAGCTGGATGCTCCGAATCATCCGCAACACCCGAGAGCGGTTGAA 2184  
Db 665 CCGCGCGGACCCACCGGTCCCGCGGATCCCGCGGTTCGCCCGCGTCCCGCGCATGGT 724  
QY 2185 ACTGCGCGCGCGCGCGCGACAGATATCGCGGCGATCCCGCGCGTACGCGCGCAATTTCCG 2244  
Db 725 GCTGCTGAAGCCGTTAGCGCGCGGTTCGCGSGGTTCGCGCGGTGGCGCCNTGGCGCGCG 784  
QY 2245 GCAGCGCGCGCGGTACAGCATGCGAATGCGCGCGCGGCGGTGTAGCGCATCTTCAACATCTC 2304  
Db 785 CCGCGCGGTTCGCGTACAGCCACCCCGCGGTTCGCGCGCATTCGCCCGCATTCGCCCATTCGC 844  
QY 2305 GCGCTACCGTCTATGCCGACAGTACCGCGCGCGCATCGCGATATG 2349  
Db 845 GCGGTTGCCGCCATTCGCCCGGTTCGCCCGCGCACCGCGGNTTG 889  
RESULT 32  
AAT91413  
ID AAT91413 standard; DNA; 913 BP.  
XX AC AAT91413;  
XX DT 12-JAN-1998 (first entry)  
XX DE Mycobacterium tuberculosis antigen Tbra32 encoding DNA.  
XX KW Antigen; immunogen; vaccine; tuberculosis; non specific adjuvant;  
XX KW skin testing; M.tuberculosis; ss.  
XX OS Mycobacterium tuberculosis.  
XX FH Key Location/Qualifiers  
XX CDS 3..208  
XX FT /\*tag= a  
XX FT /product= Antigen\_Tbra32  
XX FT /transl\_except= (pos:126..127,aa:Pro)  
XX FT /note= "Pro coded for by cc; no stop codon given"  
XX PN WO9709429-A2.  
XX PD 13-MAR-1997.  
XX PF 30-AUG-1996; 96WO-US14675.  
XX PR 12-JUL-1996; 96US-0680573.  
XX PR 01-SEP-1995; 95US-0523435.  
XX PR 22-SEP-1995; 95US-0532136.  
XX PR 22-MAR-1996; 96US-0620280.  
XX PR 05-JUN-1996; 96US-0658800.  
XX PA (CORI-) CORIXA CORP.  
XX PI Campos-neto A, Dillon DC, Houghton R, Reed SG, Skeiky YAW;  
XX PI Twardzik DR, Vedwick TH;  
XX DR WPI: 1997-192904/17.  
XX DR P-PSDB; AAW32366.  
XX PT New immunogenic polypeptide(s) from soluble M. tuberculosis antigens  
XX PT - useful for diagnosis of M. tuberculosis infection  
XX PS Claim 3; Page 64-65; 190pp; English.  
XX CC A new immunogenic polypeptide has been developed comprising an  
XX CC immunogenic part of a soluble Mycobacterium tuberculosis antigen (or









```
DT 15-JAN-2002 (first entry)
XX Mycobacterium tuberculosis strain H37Rv genome SEQ ID NO 1.
DE Mycobacterium tuberculosis; strain H37Rv; strain CDC 1551; genome;
KW variation; epidemiology; patient treatment; epidemic monitoring; ds.
XX Mycobacterium tuberculosis.
OS US6294328-B1.
XX
XX 25-SEP-2001.
XX
XX 24-JUN-1998; 98US-0103840.
XX
XX 24-JUN-1998; 98US-0103840.
XX
XX (GENO-) INST GENOMIC RES.
XX
XX Fleischmann RD, White OR, Fraser CM, Venter JC;
PI WPI; 2001-647261/74.
XX
XX Evaluating strain variation of Mycobacterium tuberculosis, comprises
PT determining the nucleotide sequence of the strain at positions in the
PT genome corresponding to positions where M. tuberculosis strains CDC
PT 1551 and H37Rv differ.
XX
XX Claim 3; SEQ ID NO 1; 3pp + Sequence Listing; English.
XX
XX The invention relates to evaluating strain variation within and between
CC different populations of the tuberculosis bacterial pathogen,
CC Mycobacterium tuberculosis or related Mycobacterium by determining the
CC nucleotide sequence of the first strain at positions in the complete
CC sequence of the genome that correspond to positions that differ in the
CC nucleotide sequences of M. tuberculosis strains CDC 1551 (AA199683) and
CC H37Rv (AA199682). The method is useful for evaluating strain variation of
CC M. tuberculosis and has valuable application in the fields of
CC tuberculosis genetics, epidemiology, patient treatment and epidemic
CC monitoring.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html?docID=6294328B1.
XX
XX Sequence 4411529 BP; 758565 A; 1449983 C; 1444602 G; 758379 T; 0 other;
SQ
Query Match 1.5%; Score 49.6; DB 22; Length 4411529;
Best Local Similarity 44.1%; Pred. No. 0.23;
Matches 208; Conservative 0; Mismatches 264; Indels 0; Gaps 0;
QY 2044 GCGGCACGGACTGCTTCGGCAGCGGCACATTCGCGCGCCGCGTCTGAAACACGCGGTA 2103
DB 336867 GCGACCCCGCGCGTCCGCGACATTTAGGCGCTGCGCCCGTCCGCGCGCGCGCC 336926
QY 2104 GAACAGGGGCGGCAATCTGGAACCTGATGCTGGAACCTGATGCTCCGAATCATCC 2163
DB 336927 GGAGGCGCGCGAGGAGTAGCGCGGTGCGCGCGTGTGCGCGCGAAGGT 336986
QY 2164 GCAACACCGGAGAGGTGTGAACTGCGCGCGCGCGACACATATGCGCGCATCCG 2223
DB 336987 GCGCGCGCTGCTCCGCCACGACGCGCGAGTGCAGCGCGCGCGCGCGCGCGCG 337046
QY 2224 CCTACGGCGCAACTTTTCGCGCAGCGGACGCGTACAGCATCGGAATCGCGCGCGGT 2283
DB 337047 CCGCGCGCGCGCGTCCGCGCGCGAGCGTGC CGCGCGCGTCCGCGCGCGCGAGAG 337106
QY 2284 GTACGATCTTCAACAGTCTCGCGGTACCGTCTATGCGGACAGTACCGCGCGCGCATGCC 2343
DB 337107 CAATCGCTTCCCGCGTCCCGCCATTCGCGCAACCGCGCGCGCGCGCGCGCG 337166
QY 2344 GATATCGAGGAGCGCGGTGAAGCGGTATCGGACGGGTGGAACACACGCTACGGGT 2403
DB 337167 GTTGGGGAACAGCCACCGGTACCAACCGGTCCGCGCGGTGCGCGCGCGCGATAAGTT 337226
```

QY 2404 CTGCGGCTCATCGCGCAACACCAACAGGACGGTGGAACTGGGAACAGCGCGGTGTTGAA 2463  
DB 337227 TTGGGAGAGCGCGCTGCGCGCGCGGTCTCTGCGGACCGAGGACAGCGCGGTACCC 337286  
QY 2464 GCGAAATGCGCGCGAGTACCCAAACCGTTCGCGCATTCGCGCGAAACCGCGCG 2515  
DB 337287 GCGCGCGCGCGCGCGCGGTGTCAGGCGCAACCGCGCGGTGCGCGCG 337338

RESULT 39  
ABK74700  
ID ABK74700 standard; DNA; 750 BP.  
XX  
AC ABK74700;  
XX  
DT 13-AUG-2002 (first entry)  
XX  
DE Bacillus licheniformis genomic sequence tag (GST) #1991.  
XX  
KW Differential gene expression; genomic sequenced tag; GST;  
KW altered culture condition; environmental stress;  
KW physiological provocation; ds.  
XX  
OS Bacillus licheniformis.  
XX  
PN WO200229113-A2.  
XX  
PD 11-APR-2002.  
XX  
PF 05-OCT-2001; 2001WO-US31437.  
XX  
PR 06-OCT-2000; 2000US-0680598.  
XX  
PR 27-MAR-2001; 2001US-279526P.  
XX  
PA (NOVO ) NOVOZYMES BIOTECH INC.  
XX  
PI Berka R, Clausen IG;  
XX  
DR WPI; 2002-416684/44.  
XX  
PT Monitoring differential expression of several genes in first Bacillus  
PT cell relative to expression of same genes in one or more second  
PT Bacillus cells by using substrate containing Bacillus genomic  
PT sequenced tag array -  
XX  
PS Claim 4; SEQ ID NO 1991; 200pp; English.  
XX  
CC The invention describes a method of monitoring differential expression of  
CC genes in a first Bacillus cell relative to expression of the genes in  
CC other Bacillus cells, comprising hybridising labelled nucleic acid probes  
CC isolated from Bacillus cells to a substrate containing array of Bacillus  
CC genomic sequenced tags (GST), examining the array, and determining  
CC relative gene expression by an observed hybridisation reporter signal of  
CC a spot in the array. The method is useful for measuring the expression of  
CC genes in a first Bacillus cell relative to expression of the same genes  
CC in one or more second Bacillus cells. The method is useful for monitoring  
CC global expression of several genes from a Bacillus cell, discovering new  
CC genes, identifying possible functions of unknown open reading frames and  
CC monitoring gene copy number variation and stability. Monitoring changes  
CC in expression of genes may be used to provide a representation of the way  
CC in which Bacillus cells adapt to changes in culture conditions, like  
CC environmental stress or other physiological provocation. Extensive  
CC follow-up characterisation is unnecessary, when one spot on an array  
CC equals one gene or one open reading frame, since sequence information is  
CC available. This sequence represents a genomic sequence tag (GST) used in  
CC the method of the invention.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

SQ Sequence 750 BP; 228 A; 165 C; 198 G; 159 T; 0 other;
Query Match 1.5%; Score 48.4; DB 24; Length 750;
Best Local Similarity 58.2%; Pred. No. 0.0061;
Matches 85; Conservative 0; Mismatches 61; Indels 0; Gaps 0;
QY 1445 CGAAAGGTACATCCGATATGTGCTACTCTCCGTAACGACATTTTCAGCAGCGGGCGGC 1504
Db 233 CAAAAGCGGCTTCATACAGCGATCGCTGGCGCAACGACATCTCCGCGCACCGGAAAGC 292
QY 1505 TGATCAAAAAGCGCGCAGCCCACTGCAACTGCAGCGGCAACACACCTATACGGGCAAAA 1564
Db 293 TGACCAAAAAGGAGACAGCGCTTTGAAGCTGGAAGGGGATAATACATATTCGCGCGGTA 352
QY 1565 CCATTATCGAAGCGGCTTCGCTGGTG 1590
Db 353 CACGGATTGATCAAGGACACTTGAG 378

Search completed: January 27, 2003, 18:45:56
Job time : 7502 secs

RESULT 40
AAI99683/c
ID AAI99683 standard; DNA; 4403765 BP.
XX
AC AAI99683;
XX
DT 15-JAN-2002 (first entry)
XX
DE Mycobacterium tuberculosis strain H37Rv genome SEQ ID NO 2.
XX
KW Mycobacterium tuberculosis; strain H37Rv; strain CDC 1551; genome;
KW variation; epidemiology; patient treatment; epidemic monitoring; ds.
XX
OS Mycobacterium tuberculosis.
XX
PN US6294328-B1.
XX
PD 25-SEP-2001.
XX
PF 24-JUN-1998; 98US-0103840.
XX
PR 24-JUN-1998; 98US-0103840.
XX
PA (GENO-) INST GENOMIC RES.
XX
PI Fleischmann RD, White OR, Fraser CM, Venter JC;
XX
WPI; 2001-647261/74.
XX
PT Evaluating strain variation of Mycobacterium tuberculosis, comprises
PT determining the nucleotide sequence of the strain at positions in the
PT genome corresponding to positions where M. tuberculosis strains CDC
PT 1551 and H37Rv differ -
XX

Claim 4; SEQ ID NO 2; 3pp + Sequence Listing; English.
XX
XX The invention relates to evaluating strain variation within and between
XX different populations of the tuberculosis bacterial pathogen,
CC Mycobacterium tuberculosis or related Mycobacterium by determining the
CC nucleotide sequence of the first strain at positions in the complete
CC sequence of the genome that correspond to positions that differ in the
CC nucleotide sequences of M. tuberculosis strains CDC 1551 (AAI99683) and
CC H37Rv (AAI99682). The method is useful for evaluating strain variation of
CC M. tuberculosis and has valuable application in the fields of
CC tuberculosis genetics, epidemiology, patient treatment and epidemic
CC monitoring.
CC
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html?DocID=6294328B1.
XX
SQ Sequence 4403765 BP; 757105 A; 1447799 C; 1441301 G; 757371 T; 189 other;
Query Match 1.4%; Score 45.2; DB 22; Length 4403765;
Best Local Similarity 57.3%; Pred. No. 3.9;

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 27, 2003, 19:00:49 ; Search time 105 Seconds  
(without alignments)  
13708.995 Million cell updates/sec

Title: US-09-830-433A-7  
Perfect score: 3204  
Sequence: 1 atgcgaacgaccccaacctt.....gcgtaggctaccggtcttga 3204

Scoring table: IDENTITY\_NUC  
Gapop 10\_0 , Gapext 1.0

Searched: 396772 seqs, 224632407 residues  
Total number of hits satisfying chosen parameters: 793544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : Published\_Applications\_NA:\*
- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
  - 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
  - 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
  - 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
  - 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
  - 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
  - 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
  - 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
  - 9: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
  - 10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
  - 11: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
  - 12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
  - 13: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
  - 14: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	270.8	8.5	286	9	US-09-928-457-95
2	48.4	1.5	750	10	US-09-974-300-1991
3	41	1.3	1866	9	US-09-738-626-1317
C 4	41	1.3	3309400	9	US-09-738-626-1
C 5	38.4	1.2	1854	9	US-09-894-844-103
6	37	1.2	1041	10	US-09-815-242-7885
C 7	37	1.2	6483	10	US-09-880-107-3699
C 8	36.8	1.1	1602	10	US-09-833-790-417
C 9	36.4	1.1	2858	9	US-09-991-496-131
10	36.4	1.1	4074	10	US-09-815-242-7899
11	35.8	1.1	3891	9	US-10-108-605-348
12	35.6	1.1	4689	9	US-09-860-846-34
13	35.6	1.1	4689	10	US-09-861-289-34
14	35.6	1.1	36778	9	US-09-860-846-5
C 15	35.6	1.1	36778	10	US-09-861-289-5
C 16	35.4	1.1	453	10	US-09-960-352-11913
C 17	35.4	1.1	935	10	US-09-770-445-373
18	35.2	1.1	213	10	US-09-864-761-21194
C 19	35.2	1.1	422	10	US-09-924-035A-476

C 20	35.2	1.1	655	10	US-09-770-149-502	Sequence 502, App
21	35.2	1.1	1693	10	US-09-867-550-1669	Sequence 1669, Ap
22	35	1.1	418	10	US-09-960-352-11256	Sequence 11256, A
23	35	1.1	1761	10	US-09-969-347-2	Sequence 2, Appli
24	35	1.1	1761	10	US-09-880-107-3665	Sequence 3665, Ap
25	35	1.1	1904	10	US-09-925-301-594	Sequence 594, App
26	35	1.1	3126	10	US-09-815-242-4020	Sequence 4020, Ap
C 27	35	1.1	88421	9	US-09-976-059-1	Sequence 1, Appli
28	34.6	1.1	1296	10	US-09-815-242-4090	Sequence 4090, Ap
29	34.4	1.1	3786	10	US-09-815-242-7865	Sequence 7865, Ap
C 30	34.2	1.1	756	9	US-09-938-842A-2617	Sequence 2617, Ap
31	34.2	1.1	1294	10	US-09-850-964-4	Sequence 4, Appli
32	34.2	1.1	1488	10	US-09-815-242-7862	Sequence 7862, Ap
33	34.2	1.1	2100	9	US-09-738-626-1740	Sequence 1740, Ap
34	34.2	1.1	2289	10	US-09-850-964-3	Sequence 3, Appli
35	34.2	1.1	3042	9	US-09-712-363-68	Sequence 68, Appli
36	34.2	1.1	3309400	9	US-09-738-626-1	Sequence 1, Appli
37	34	1.1	644	10	US-09-867-550-875	Sequence 875, App
C 38	34	1.1	2936	9	US-10-108-605-54	Sequence 54, Appl
C 39	34	1.1	9210	9	US-09-712-363-100	Sequence 100, App
40	33.8	1.1	969	10	US-09-815-242-7933	Sequence 7933, Ap
41	33.8	1.1	1768	10	US-09-833-381-523	Sequence 523, App
42	33.8	1.1	2089	9	US-10-051-902-25	Sequence 25, Appl
43	33.8	1.1	2089	9	US-10-051-909-25	Sequence 25, Appl
44	33.8	1.1	2340	10	US-09-815-242-7995	Sequence 7995, Ap
C 45	33.8	1.1	7414	10	US-09-764-853-876	Sequence 876, App

ALIGNMENTS

RESULT 1  
US-09-928-457-95  
; Sequence 95, Application US/09928457  
; Patent No. US20020164603A1  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: DNA, specific proteins and peptides  
; TITLE OF INVENTION: of the Neisseria meningitidis species bacteria, method  
; TITLE OF INVENTION: for obtaining them and their biological application.  
; NUMBER OF SEQUENCES: 99  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (OEB)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/928,457  
; FILING DATE: 2001-08-14  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/214,759  
; FILING DATE: 199-12-10  
; INFORMATION FOR SEQ ID NO: 95:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 286 base pairs  
; TYPE: nucleotide  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
US-09-928-457-95

Query Match 8.5%; Score 270.8; DB 9; Length 286;  
Best Local Similarity 99.0%; Pred. No. 3.8e-74;  
Matches 283; Conservative 0; Mismatches 2; Indels 1; Gaps 1;  
QY 865 AATTCGGAGGACGATACCGCAAGCGTTCGCCCTATTTCGGCGGTGATAAACAGAC 924  
Db 1 AATTCGGAGGACGATACCGCAAGCGTTCGCCCTATTTCGGCGGTGATAAACAGAC 60  
QY 925 GAGGGTATCCGGCTGATGCAACAGACGCGATTACGGCAACTTCCTACCACTATCGGTAAT 984  
|||||





RESULT 10  
US-09-815-242-7899  
; Sequence 7899, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert



```
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA 011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7899
; LENGTH: 4074
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(4074)
US-09-815-242-7899

Query Match 1.1%; Score 36.4; DB 10; Length 4074;
Best Local Similarity 59.8%; Pred. No. 1.8;
Matches 61; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 1477 CGTACGACATTTTCAGCGAGCGCGCTGATCAAAAAGCGCGACGCCAACTGCAACTG 1536
Db 1246 CGTACGAGATCAAGGTCGCGCTCTGAGCAAGGAAGACATCATCGATGTGCTCAAG 1305

QY 1537 CACGGCAACAACACCTATACGGGCAAAACCATTTATCGAAGGC 1578
Db 1306 ACCCTCGTCAGATCGGTAAACGGCAAGGGCATCGTCGATGAC 1347

RESULT 11
US-10-108-605-348
; Sequence 348, Application US/10108605
; Patent No. US20020160934A1
; GENERAL INFORMATION:
; APPLICANT: Broadus, Julie
; APPLICANT: Stam, Lynn
; APPLICANT: Bachmann, Jane
; APPLICANT: Kamdar, Kim
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCODE
; FILE REFERENCE: 31133B
; CURRENT APPLICATION NUMBER: US/10/108,605
; PRIOR FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: US 09/761,142
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/176,418
; PRIOR FILING DATE: 2000-01-14
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 348
; LENGTH: 3891
; TYPE: DNA
```

```
; ORGANISM: Drosophila melanogaster
US-10-108-605-348

Query Match 1.1%; Score 35.8; DB 9; Length 3891;
Best Local Similarity 49.2%; Pred. No. 2.7;
Matches 94; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

QY 2147 ATGCTCCGAATCATCCGCAACACCCGAGAGCGGTTGAAACTGCGCGCGACCGACAG 2206
Db 3223 ACGTGCACCTCATCCGCACTGGCGCAGCAGCAAGCTGCTCCGATCCGACACAG 3282
QY 2207 ATATGCCGGGCATCCGCCCTTACGGCGCAACTTTCCGCGCAGCGCGACCGCTACAGCATG 2266
Db 3283 TTGCGCGCGCAGTCCCAACCGCTAGCGGTCCCAACCGCGCGCAGCAGCGCTGCTCCCGCG 3342
QY 2267 CGAATGCCGCGCGAGGTGTAGGCATCTTCAACAGTCTTCGCGCGTACCGTCTATGCCGACA 2326
Db 3343 CACGTAGTGTGTCGCCCAACGGTGTCTCCCTTTTCAGCGCGCACCGATTTCGCTCACGGGA 3402
QY 2327 GTACCGCGCGC 2337
Db 3403 GGACTCCCGCC 3413

RESULT 12
US-09-860-846-34
; Sequence 34, Application US/09860846
; Patent No. US20020164742A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/860,846
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 34
; LENGTH: 4689
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-860-846-34

Query Match 1.1%; Score 35.6; DB 9; Length 4689;
Best Local Similarity 45.4%; Pred. No. 3.6;
Matches 128; Conservative 0; Mismatches 154; Indels 0; Gaps 0;

QY 2077 GCGCGCGCGCGTGTGAAACACGCGGTAGAACAGGGCGGAGCAATCTGGAAACCTGATG 2136
Db 3406 GCCCACCAGCGAGCGCAACACCTCTCTCTGTCAGCGCGCAGCGCAACAGCCCGCGGA 3465
QY 2137 GTGCACTGGATGCTCCGAATCATCCGCAACACCCGAGAGCGGTTGAACTGCGCGCGCC 2196
Db 3466 GCCACCAACTCACCGCGCACTACCGCATCGGGCGCGCCGCGCTACCATCGCGCGCTGC 3525
QY 2197 GACCGCACAGATATGCCGGGCACTCCGCCCTTACGGCGCAACTTTCCGCGCAGCGCGACGC 2256
Db 3526 GACGTGCCGACCCCGCAGCGCATCGGCAACCTCTCTCAGCGCCATCCCGCGCAGACGCC 3585
QY 2257 GTACAGCATCGAATGCCGCGCGAGCGGTGTACGCACTTTCAACAGTCTTCGCGCGTACCGTC 2316
Db 3586 CTCACCGCGCGTCTGTCACACACCGCGCGCAGCGCGCGCATCGGCTGGAGTCAACCGGC 3645
QY 2317 TATGCCGACAGTACCGCGCGCCCATTCGCGATATGACGAGGAGCG 2358
Db 3646 CCGGAGGACATCGCGCGCATCTCTGGCGCGGAAGACGAGCGGC 3687

RESULT 13
```



```
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 11913
; LENGTH: 453
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 51-LIB3058-002-Q1-K1-E4
US-09-960-352-11913

Query Match      1.1%; Score 35.4; DB 10; Length 453;
Best Local Similarity 48.3%; Pred. No. 0.87;
Matches 99; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

QY 2137 GTGCACTGGATGCCCTCCGAATCATCCGCAACACCGAGACGGTTGAAACTGCGGCGCC 2196
      ||||| | | ||||| | | ||||| | | ||||| | | ||||| | | ||||| | |
Db 396 GTGCTAGCGCTGCTCCGCGACTTGTGTACCGCGCGCGCGCGTGCACCGCGGTAGTC 337

QY 2197 GACCGCACAGATATCGCGGCATCCGCCCTACGGCGCAACTTTCGCGCGAGCGGCAGCC 2256
      ||||| | | ||||| | | ||||| | | ||||| | | ||||| | | ||||| | |
Db 336 GTCGCGCGCGCTGCGGCGCTAGTGTCTCGCGCGCGGACTTGTGTAGCCACCGCGCC 277

QY 2257 GTACAGCATGCGAATGCCGCGGAGTGTACGCATCTTCAACAGTCTCGGCGCTACCGTC 2316
      ||||| | | ||||| | | ||||| | | ||||| | | ||||| | | ||||| | |
Db 276 GGTGCCACGCGCTAGTGTGCGCGCGGACCTGCGTACCCATCGCGCGCGCTGCGGCC 217

QY 2317 TATGCCGACAGTACCGCGCGCCCATG 2341
      ||||| | ||||| | ||
Db 216 GTAGTGTGTCATCGCGCGCGACTTG 192

RESULT 17
US-09-770-445-373
; Sequence 373, Application US/09770445
; Patent No. US20020023281A1
; GENERAL INFORMATION:
; APPLICANT: Goriach, Jörn
; APPLICANT: An, Yong-Qiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.
; APPLICANT: Raines, Tracy M.
; APPLICANT: Yu, Yang
; APPLICANT: Rameaka, Joshua G.
; APPLICANT: Page, Amy
; APPLICANT: Matthew, Abraham V.
; APPLICANT: Ledford, Brooke L.
; APPLICANT: Woessner, Jeffrey P.
; APPLICANT: Haas, William David
; APPLICANT: Garcia, Carlos A.
; APPLICANT: Krieker, Maja
; APPLICANT: Slader, Ted
; APPLICANT: Davis, Keith R.
; APPLICANT: Allen, Keith
; APPLICANT: Hoffman, Neil
; APPLICANT: Hurban, Patrick
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; FILE REFERENCE: 2023US (PARA-012PRV)
; CURRENT APPLICATION NUMBER: US/09/770,445
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: US 60/178,472
; PRIOR FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 999
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 373
; LENGTH: 935
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-770-445-373

Query Match      1.1%; Score 35.4; DB 10; Length 935;
Best Local Similarity 45.7%; Pred. No. 1.4;
Matches 123; Conservative 0; Mismatches 146; Indels 0; Gaps 0;

QY 2616 ACGGACGATGCGGGCGGATATCGGCTATCTCAAAGGCGCTTCTCTTACGAGCGCTACAA 2675
      ||||| | | ||||| | | ||||| | | ||||| | | ||||| | | ||||| | |
Db 516 ATGTCATCACGATCTCGAAACGAAAGCGTTCAACACTCTCTACTCCACGAGAGACCAA 575

QY 2676 AAACAGCATCAGCGCGACCGTGGGCGGAGCAACATCGCGAAGCGCGCTCAACGGCAC 2735
      ||||| | | ||||| | | ||||| | | ||||| | | ||||| | | ||||| | |
Db 576 AAGCGGCGACACGCCAACGCCGCTGGCGAAATTAAGGATGAGGCAAGCGCACGAGGC 635

QY 2736 GCTGATGCAGCTGGGCGGCGTGTCAACGTTTCCCGTTTCCGCAACGGGAGATTT 2795
      ||||| | | ||||| | | ||||| | | ||||| | | ||||| | | ||||| | |
Db 636 CGAGAAGCTGAGCAATCCAGTAGAGGATACCCACGGAGGAGTAGTGTACCACCGATGA 695

QY 2796 GACGGTCGAAGCGGCTGCGCTACGACCTGTCTCAACAGAGATGCATTCGCCGAAAAAGG 2855
      ||||| | | ||||| | | ||||| | | ||||| | | ||||| | | ||||| | |
Db 696 AAGCAGCGAAAGTGAGCGGCAAGGTTAACGTTGTCACCGAGAGATGTTGGCACCACCTGAGA 755

QY 2856 CAGTGTCTTTGGGCTGGAGCGGCAACGCC 2884
      ||||| | | ||||| ||||
Db 756 CAGCGACGAAGAGTCCAAGGCGATGAGCC 784

RESULT 18
US-09-864-761-21194
; Sequence 21194, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
```





RESULT 26

US-09-815-242-4020

; Sequence 4020, Application US/09815242

; Patent No. US20020061569A1

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari L.

```

RESULT 27
US-09-976-059-1/c
; Sequence 1, Application US/09976059
; Patent No. US2002016474A1
; GENERAL INFORMATION:
; APPLICANT: Farnet, Chris
; APPLICANT: Zazopoulos, Emmanuel
; APPLICANT: Staifa, Alfredo
; TITLE OF INVENTION: Genes and Proteins for Biosynthesis of Ramoplanin
; FILE REFERENCE: 3019-PCT
; CURRENT APPLICATION NUMBER: US/09/976.059
; CURRENT FILING DATE: 2001-10-15
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 88421
; TYPE: DNA
; ORGANISM: Actinoplanes sp.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2077)..(3078)
; OTHER INFORMATION: ORF 1; positive strandedness
; NAME/KEY: misc_feature
; LOCATION: (3118)..(4032)
; OTHER INFORMATION: ORF 2; positive strandedness
; NAME/KEY: misc_feature
; LOCATION: (4038)..(5048)

```



```
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4090
; LENGTH: 1296
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-4090

Query Match      1.1%  Score 34.6;  DB 10;  Length 1296;
Best Local Similarity 48.7%;  Pred. No. 3.1;
Matches 94;  Conservative 0;  Mismatches 99;  Indels 0;  Gaps 0;

QY 2626 GCGGCGGATATCTCTCAAGGCTGTCTCTACGGACGCTACAAAAACAGCATC 2685
      ||||| || ||||| ||||| || ||||| || ||||| || ||||| || |||||
Db 346 GCGTGGCATCGCTACACCTCGGCTGGTTTCGAGAAGGAGCGGCGACCGCGATG 405

QY 2686 AGCCGACGACCGGTGGGAGACATCGCGAAGGACGCTCAACGGACGCTGATGAG 2745
      || || || || || || || || || || || || || || || || || || ||
Db 406 GGCATCTTGGCGCGGCAACGCGCGCGGCGGCGGATACCAACCTGGTGGCGCGATGATC 465

QY 2746 CTGGGCGCACTGGCGGTGTCAACGTTCCTTCCTCCGCAACGGGAGATTTCACGGTCA 2805
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 466 GTGGTGGCTTTCGGCTGGCGGATGTTGCCGAGGTCTACTCGGTGGCCATCTGTTGTCAC 525

QY 2806 GCGGTCTGCGCT 2818
      || || || || ||
Db 526 GCGTGGCTGTTCT 538

RESULT 29
US-09-815-242-7865
; Sequence 7865, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815.242
; CURRENT FILING DATE: 2001-03-21
; PRIOR FILING DATE: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7865
; LENGTH: 3786
; TYPE: DNA
```

```
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(3786)
US-09-815-242-7865

Query Match      1.1%  Score 34.4;  DB 10;  Length 3786;
Best Local Similarity 54.8%;  Pred. No. 7.3;
Matches 68;  Conservative 0;  Mismatches 56;  Indels 0;  Gaps 0;

QY 44 CTGCCATGGCGGTAGCTTGTGCAACACACTTTCTGCTGCTTAGGCGGCGCGCGCA 103
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1713 CGGCCAAAGCGGCGCGCTGGCGCACTATGTGGCCAGGAGAGCTGCGCCGCGAGAC 1772

QY 104 CTTCTGCGCGGACTTCAATGCGGCGGCAACGCTATCGGCAACAGCAGACGACAA 163
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1773 CGGTGGCGCGCTGCGCTTCGGCTTGGACTGGAGCGGCGCGCGGCGGAGATGAACGG 1832

QY 164 CAGC 167
      || |
Db 1833 CACC 1836

RESULT 30
US-09-938-842A-2617/c
; Sequence 2617, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR FILING DATE: 60/227,866
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2617
; LENGTH: 756
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-2617

Query Match      1.1%  Score 34.2;  DB 9;  Length 756;
Best Local Similarity 45.4%;  Pred. No. 2.9;  Mismatches 148;  Indels 0;  Gaps 0;
Matches 123;  Conservative 0;

QY 2616 ACGGCACGATCGGCGCGATATCGGCTATCTCAAGGCGCTTCTCTCTACGACGCTACAA 2675
      || || || || || || || || || || || || || || || || || || ||
Db 457 ATGTCATCAGCATCTCGAAACGAAAGGTTCAACACTCTCTCTCTCTCTCTCTCTCTCT 398

QY 2676 AAACAGCATACGCGCGACCGGTCGGGAGACATCGGGAAGCAGCGCTCAACGGCAC 2735
      || || || || || || || || || || || || || || || || || || ||
Db 397 AAGCGCGCACACCAAGCCACCGGTTAAAGGATGAGGCAACGCGACGCGGAGC 338

QY 2736 GCTGATGAGCTGGGCGCACTGGGCGGTGTCAAGCTTTCGCGCAACGGGAGATT 2795
      || || || || || || || || || || || || || || || || || || ||
Db 337 CGAAGAGCTGAGCAATCCAGTAGAGATACCGGAGAGAGATGATGTTACCACTGA 278

QY 2796 GACGGTCGAAGCGGCTCTGCGCTACGACTCTCAACAGGATGATCTCGCGGAAAAAGG 2855
      || || || || || || || || || || || || || || || || || || ||
Db 277 AAGCACCAGAAAGTAGGCGCAGGGTTAAGCTGTCCACCAAGAGATGTTGGCACAACACTGAGA 218

QY 2856 CAGTGTCTTTGGGCTGGAGCGGCAACAGCGCTC 2886
      || || || || || || || || || || || || || || || || || || ||
Db 217 CAGCGACGAGAGTCCAAAGGCATGAGCCAC 187
```



```

: PRIOR APPLICATION NUMBER: 60/207,727
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: 60/242,578
: PRIOR FILING DATE: 2000-10-23
: PRIOR APPLICATION NUMBER: 60/253,625
: PRIOR FILING DATE: 2000-11-27
: PRIOR APPLICATION NUMBER: 60/257,931
: PRIOR FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: 60/269,308
: PRIOR FILING DATE: 2001-02-16
: NUMBER OF SEQ ID NOS: 14110
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 7862
: LENGTH: 1488
: TYPE: DNA
: ORGANISM: Pseudomonas aeruginosa
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1)...(1488)
US-09-815-242-7862

Query Match          1.1%   Score 34.2; DB 10; Length 1488;
Best Local Similarity 56.8%; Pred. No. 4.5;
Matches 63; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 2652 CCTGTTCTCTACGGACGCTACAAAACAGCATCAGCGCCAGCAGCCGTGCGGACGACA 2711
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 51  CCTGGTGCTCGCGGTAGCGGAAGCCGCAAGCTCGGCGCCAGCCGCAAGGGCGGTGACGA 110

QY 2712 TGGCAGGAGCGAGCTCAACGGCAGCGTGTATGCACGTGGGCGGCATGGGCGG 2762
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 111 GCGCACCGCGCGCGCCATCAGCGCAGTGCTCAAGCGTGGCGACCTGGCGCGG 161

RESULT 33
US-09-738-626-1740
: Sequence 1740, Application US/09738626
: Publication No. US20020197605A1
: GENERAL INFORMATION:
: APPLICANT: NAKAGAWA, SATOSHI
: APPLICANT: MIZOGUCHI, HIROSHI
: APPLICANT: ANDO, SEIKO
: APPLICANT: HAYASHI, MIKIRO
: APPLICANT: OCHIAI, KEIKO
: APPLICANT: YOKOI, HARUHIKO
: APPLICANT: TATEISHI, NAKO
: APPLICANT: SENOH, AKIHIRO
: APPLICANT: IKEDA, MASATO
: APPLICANT: OZAKI, AKIO
: TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
: FILE REFERENCE: 249-125
: CURRENT APPLICATION NUMBER: US/09/738,626
: CURRENT FILING DATE: 2000-12-18
: PRIOR APPLICATION NUMBER: JP 99/377484
: PRIOR FILING DATE: 1999-12-16
: PRIOR APPLICATION NUMBER: JP 00/159162
: PRIOR FILING DATE: 2000-04-07
: PRIOR APPLICATION NUMBER: JP 00/280988
: PRIOR FILING DATE: 2000-08-03
: NUMBER OF SEQ ID NOS: 7059
: SOFTWARE: PatentIn ver. 3.0
: SEQ ID NO 1740
: LENGTH: 2100
: TYPE: DNA
: ORGANISM: Corynebacterium glutamicum
US-09-738-626-1740

Query Match          1.1%   Score 34.2; DB 9; Length 2100;
Best Local Similarity 46.8%; Pred. No. 5.7;
Matches 108; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

QY 2216 GCATCCGCCCTACGGCGCAACTTTCCGCGCAGCGCGGCGGTACAGCATGCGAATGCCG 2275

```

```
Db 1307 GTAACCTGCACCTTCGGTATCCGTGAGCAGCGTATGGGATCCATCTCAACGGCATTTCCC 1366
QY 2276 CCGAGGGTGTAGCATCTTCAACAGTCTCGCGGTACCGTCTATCCGACAGTACCGCGC 2335
Db 1367 TCCACGGTGGCACCAGCCCATACGCGGAACCTTCTCATCTTCTCCGACTACATGCGTC 1426
QY 2336 CCCATGCCGATATGACGAGCCCGCGCTGAAGACCGTATCGGACGGTTGGACCAACG 2395
Db 1427 CTGCAGTTTCGCTTTCAGCTCTCATGAGACGAGCGCTTACTACGCTTGGACCCAGACT 1486
QY 2396 CTACGGGTCTGCGGCTCATCGCGCAAAACCAACAGGACGGTGAACGTTGG 2446
Db 1487 CCATCGGTCTGGCGCAAGATGCCCAACCCACCGCTGTGAACCTTGG 1537

RESULT 34
US-09-850-964-3
; Sequence 3, Application US/09850964
; Patent No. US20020104121A1
; GENERAL INFORMATION:
; APPLICANT: MCELROY, DAVID
; APPLICANT: KRIZ, ALAN L.
; APPLICANT: OROZCO, EMIL
; APPLICANT: GRIFFOR, MATT
; TITLE OF INVENTION: THE MAIZE A3 PROMOTER AND RICE ACTIN 2 INTRON AND
; FILE OF INVENTION: METHODS FOR USE THEREOF
; FILE REFERENCE: DEKM:166
; CURRENT APPLICATION NUMBER: US/09/850,964
; PRIOR FILING DATE: 2001-05-07
; PRIOR APPLICATION NUMBER: 09/312,038
; PRIOR FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2289
; TYPE: DNA
; ORGANISM: Zea mays/rice
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (889)..(1825)
; OTHER INFORMATION: N = A or C or G or T
US-09-850-964-3

Query Match 1.1%; Score 34.2; DB 10; Length 2289;
Best Local Similarity 47.1%; Pred. No. 6;
Matches 105; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

QY 2115 CAGCAATCTGAAACCTGATGGTCACTGGATGCCTCCGAATCATCCGCAACACCGCA 2174
Db 472 CATCAGGTAGCGAACAGGATACTTGGCGCGGGCGCCACCGCTCAGTTCTCTCCAGT 531
QY 2175 GACGGTTGAAACTGGCGCGCGGACAGATATGCGGGCATCGCCCGCTAGCGGCG 2234
Db 532 TCCGTCGGAGTTGCGTAGCGCGGCGCACCGTTGCCACCGCGTAGCGCGCACACAC 591
QY 2235 AACTTTCCGCGACGCGGACCGGTACAGCATGCGAATGCCCGGACGGGTGACGATCTT 2294
Db 592 CACCGTAGCCACCGCGCACCGTAGCGCGCGCCACCGCGGTAGCGCGCGCGGCAT 651
QY 2295 CAACAGTCTCCCGCTACCGTCTATGCGGACAGTACCGCGCC 2337
Db 652 CACGGCGCCGCCACCGCGTAGCGCGGCTCCACGCGCCACC 694

RESULT 35
US-09-712-363-68
; Sequence 68, Application US/09712363
; Patent No. US20020164588A1
; GENERAL INFORMATION:
; APPLICANT: Eisenberg, David
; APPLICANT: Rotstein, Sergio H.
; APPLICANT: Marcotte, Edward M.
; TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND
```

```
; TITLE OF INVENTION: INTERACTIONS OF PROTEINS BY COMPARATIVE ANALYSIS
; FILE REFERENCE: 07419-032001
; CURRENT APPLICATION NUMBER: US/09/712,363
; CURRENT FILING DATE: 2000-11-13
; PRIOR APPLICATION NUMBER: PCT/US00/02246
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/179,531
; PRIOR FILING DATE: 2000-02-01
; PRIOR APPLICATION NUMBER: 60/117,844
; PRIOR FILING DATE: 1999-01-29
; PRIOR APPLICATION NUMBER: 60/118,206,
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: 60/126,593
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 60/134,093
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/134,092
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/165,124
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/165,086
; NUMBER OF SEQ ID NOS: 292
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 68
; LENGTH: 3042
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
US-09-712-363-68

Query Match 1.1%; Score 34.2; DB 9; Length 3042;
Best Local Similarity 50.9%; Pred. No. 7.3;
Matches 81; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

QY 1231 TCCCTTTCCGACCCATCGTAACCGGACGCGGCTGTGCTGCAGAAATACCCGTGG 1290
Db 1102 TCGGGTAGCGGCGCGGTCTCGACGAGGCTGGGTTTGGCGTCTGCTGCCGTCTCGTGG 1161
QY 1291 ATGAGCAACGACAACCTGCGTACCACGCTGTGACAACGGCTCAGGACATCGGTGCAGTC 1350
Db 1162 GACCGCGCGCCAAAGCTGGCTTGGTCTGTCCGATATACCCGCGTGCAGCGGTGGTG 1221
QY 1351 GCGCTGACAGCAAGTTGCGTGGGACTGCTGGATGCG 1389
Db 1222 GGCAAGGCCACCAAGTTGCGCGCGGAGCAGCTCGTCGAG 1260

RESULT 36
US-09-738-626-1
; Sequence 1, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIALI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
```



;  
; PRIOR APPLICATION NUMBER: 60/165,086  
; PRIOR FILING DATE: 1999-11-12  
; NUMBER OF SEQ ID NOS: 292  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 100  
; LENGTH: 9210  
; TYPE: DNA  
; ORGANISM: Mycobacterium tuberculosis  
US-09-712-363-100

Query Match 1.1%; Score 34; DB 9; Length 9210;  
Best Local Similarity 44.7%; Pred. No. 17;  
Matches 176; Conservative 0; Mismatches 215; Indels 3; Gaps 1;  
QY 2704 GACGAACTCGGAGCGAGCTCAACGCGACGCTGATCAGCTGGCGGCGCTG 2763  
DB 2985 GACGGCGACGCTCAAGCGACGCTGCTGCGCGCTGCTGCGCGCGCGCGGATGA 2926  
QY 2764 GTCAAGCTTCCGTTTCCGCAACGGGAGATTTGACGGTTCGAAGCGGCTGCGCTACGAC 2823  
DB 2925 GTGTGTGGCGCGCGGCTTTTCGGGTCCATCGTGCACCTGCCATTTCGGCGGCTCGGCGAT 2866  
QY 2824 CTGCTCAACAGATGTCATTCGCGCAAAAGCAGTGTCTTTGGCTGGAGCGGCAACAGC 2883  
DB 2865 CCGATGACACCGGTTGGTTCACGGTGGCGACCGCGCCACATCGGGTGCCTCGAGGAC 2806  
QY 2884 CTCACGTGAAGGACACACTGTGTCGAGCTCGCGGCTCTGAAGCTGTCGCAACCCCTTGAGCGAT 2943  
DB 2805 GACAGCAACGGTTCGGCGCACGTCGCGCGCGCC---AGCGGCGCGACGACATCCTT 2749  
QY 2944 AAAGCGCTCTGTTTCAACGGCGGCGGTGGAACGCGACCTGAACGGACGCGACTACAG 3003  
DB 2748 CGGCTCGACACCGCGCGGAGCACTTTCATGATTGGGCTTGCTCGAAACGGTCCAGCAA 2689  
QY 3004 GTAACGGCGGCTTTACCGCGCGGCTGACGCAACCGGCAAGCGGGCGACGCAATATG 3063  
DB 2688 CTCACCGCGGTTTCATCATCCGGTGTATCGCGGTACCGACGCGGTGCCCGGAATGAT 2629  
QY 3064 CGCACACCGCGCTGTTGCGGCTCTGGCGCGG 3097  
DB 2628 GCACACCGCATCGGCTGCTAGCGGCGCTGTGG 2595

RESULT 40  
US-09-815-242-7933  
; Sequence 7933, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; FILE REFERENCE: ELTRA.011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22

;  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7933  
; LENGTH: 969  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(969)  
US-09-815-242-7933  
Query Match 1.1%; Score 33.8; DB 10; Length 969;  
Best Local Similarity 47.8%; Pred. No. 4.5;  
Matches 98; Conservative 0; Mismatches 107; Indels 0; Gaps 0;  
QY 2725 GTCAACGGCACGCTGATGACGTGGGCGGCTGCGGGGTGTCAACGTTCCGTTTGGCGCA 2784  
DB 700 GGCAAGCCCACTGCGGCTGATCGTCACTGCGCGGACCGGACCGAGGAACAGGCCGG 759  
QY 2785 ACGGGAGATTTGACGGTTCGAAGCGGCTGCGCTAGGACCTGCTCAACAGGATGCATTTC 2844  
DB 760 CTGGTACGCAAAAGCCATCCAGCAAGGTGGCAGCCAGGACCTGGAGAGCTGTGTGCCGCC 819  
QY 2845 GCCGAAAAAGGAGTGTCTTTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTC 2904  
DB 820 GTCGAGGACGCGGCGGCTGCACTATACCGCCAACTCGCCGCGACTACGCCGCCGC 879  
QY 2905 GGACTCGCGGGTCTGAAGCTGTGCG 2929  
DB 880 GCCATCGCCTGCTGGATACGCTGC 904

Search completed: January 28, 2003, 00:26:14  
Job time : 14101 secs

GenCore version 5.1.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 27, 2003, 16:40:30 ; Search time 2876 Seconds

(without alignments)

18042.538 Million cell updates/sec

Title: US-09-830-433a-7

Perfect score: 3204

Sequence: 1 atgcgaacgaccccaacctt.....gcgtaggctaccggtctga 3204

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_fod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
C 1	57.4	1.8	910	17	CNS0060N	AL065629 Drosophil
C 2	47.4	1.5	580	14	B0620456	B0620456 Tatr1156f
C 3	46.4	1.4	925	17	CNS0091P	AL053013 Drosophil
C 4	45.8	1.4	902	17	CNS006QP	AL065804 Drosophil
5	45.6	1.4	471	10	AW707761	AW707761 832012E07
6	44.8	1.4	1101	17	CNS012S8	AL101954 Drosophil

7	44.6	1.4	584	9	AL824948	AL824948
8	44	1.4	488	12	BE726719	BE726719 894093B05
9	44	1.4	490	10	AV639324	AV639324 AV6393203
10	44	1.4	497	10	AV636203	AV636203 AV636203
11	44	1.4	513	10	AV636123	AV636123 AV636123
12	44	1.4	557	14	B0822928	B0822928 1030104H1
13	44	1.4	562	10	AV387616	AV387616 AV387616
14	44	1.4	634	13	BI875064	BI875064 963121E11
15	43.8	1.4	1087	17	CNS0106Y	AL098596 Drosophil
16	43.8	1.4	1101	17	CNS017X5	AL108611 Drosophil
17	43.2	1.3	655	10	AW676029	AW676029 832001E09
18	43	1.3	477	10	AV633578	AV633578 AV633578
19	42.2	1.3	358	13	BG947100	BG947100 IPl_2_E12
20	42.2	1.3	755	10	BE041076	BE041076 OF19D09 O
C 21	42	1.3	1101	17	CNS017GS	AL108022 Drosophil
C 22	42	1.3	1101	17	CNS017SY	AL108460 Drosophil
23	41.8	1.3	536	9	AL716352	AL716352 AL716352
24	41.8	1.3	538	14	BQ283304	BQ283304 WHE3088_D
25	41.8	1.3	650	9	AL721811	AL721811 AL721811
26	41.8	1.3	766	9	AL721550	AL721550 AL721550
27	41.8	1.3	772	9	AL716995	AL716995 AL716995
28	41.8	1.3	790	9	AL716744	AL716744 AL716744
C 29	41.8	1.3	792	9	AL717085	AL717085 AL717085
30	41.8	1.3	844	17	CNS0052P	AL056652 Drosophil
31	41.6	1.3	645	17	CNS01213	AL101589 Drosophil
32	41.4	1.3	444	14	B0759294	B0759294 EBP101_SQ
33	41.4	1.3	976	12	BE778605	BE778605 601466144
34	41.2	1.3	588	13	BI778455	BI778455 EBR007_SQ
35	41.2	1.3	847	17	CNS021SP	AL177298 Tetraodon
36	41	1.3	573	12	BG368419	BG368419 HVSME1001
37	40.8	1.3	540	10	AW006888	AW006888 ws15f06.x
C 38	40.8	1.3	631	13	BI790356	BI790356 Id06e08.x
39	40.4	1.3	474	13	BM068591	BM068591 WHE3459_D
40	40.4	1.3	1600	17	AG058272	AG058272 Pan trogl
C 41	40.2	1.3	497	12	BF657250	BF657250 OV2_22_E0
42	40.2	1.3	595	9	AL716267	AL716267 AL716267
43	40.2	1.3	692	17	CNS007WH	AL050923 Drosophil
44	40.2	1.3	715	9	AL715439	AL715439 AL715439
45	40.2	1.3	722	9	AL728402	AL728402 AL728402

ALIGNMENTS

RESULT 1	CNS0060N/c	910 bp	DNA	linear	GSS 03-JUN-1999
LOCUS	Drosophila melanogaster genome survey sequence T7 end of BAC #				
DEFINITION	BACr14J21 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.				
ACCESSION	AL065629				
VERSION	AL065629.1	GI:4944698			
KEYWORDS	GSS.				
SOURCE	Drosophila melanogaster.				
ORGANISM	Drosophila melanogaster.				
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				
REFERENCE	1 (bases 1 to 910)				
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)				
COMMENT	Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <a href="http://www.fruitfly.org">http://www.fruitfly.org</a> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammeter in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial				









Db 846 SNNANNNAVSSS 858

RESULT 7  
AL824948

LOCUS AL824948 p:234 Triticum aestivum cDNA clone A06\_p234\_plate\_11, mRNA  
DEFINITION sequence.

ACCESSION AL824948 584 bp mRNA linear EST 15-JUL-2002  
VERSION AL824948  
KEYWORDS EST.  
SOURCE bread wheat.  
ORGANISM Triticum aestivum

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
AUTHORS Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;  
1 (bases 1 to 584) Triticaceae; Triticum.  
Wilson,I., Beswick,R., Shepherd,S., Barker,G., Parker,J., Owen,P.,  
Edwards,D., Coghil,J., Holdsworth,M., Lenton,J., Shewry,P. and  
Edwards,K.

TITLE A BBRC-funded wheat EST resource for the academic community  
JOURNAL Unpublished (2002)  
COMMENT Contact: Barker G  
Institute of Arable Crop Research  
Long Ashton, Bristol BS41 9AF United Kingdom.

FEATURES  
source Location/Qualifiers  
1..584  
/organism="Triticum aestivum"  
/cultivar="mercia"  
/db\_xref="taxon:4565"  
/clone="A06\_p234\_plate\_11"  
/clone\_lib="p:234"  
/tissue\_type="drought stressed seedlings"  
/dev\_stage="21 days old"  
BASE COUNT 102 a 221 c 181 g 80 t  
ORIGIN

Query Match 1.4%; Score 44.6; DB 9; Length 584;  
Best Local Similarity 56.5%; Pred. No. 0.27;  
Matches 83; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 1219 ATTCCCGGACATCCTTTTCGGACACCATCGTACCGGGACGGGGCTGCTGCTGCAG 1278  
|| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
Db 351 ATCTCCGGGACGTCTGTCGTGCGCCAGCTGAGCGGCTGCTCGCGGCG 410  
|| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
QY 1279 AAATACCCGTGGATGACACGACAACTCGGTACCACTGCTGACAAAGCGCTCAGGAC 1338  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 411 CGCACCCGGAGTGGAGCCCGGGGGGTCAAGTCGGCGCTCATGCCACGGCTCAAC 470  
|| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
QY 1339 ATCGGTGACGTGGCGGTGGACAGCAAG 1365  
|| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
Db 471 CTGGACAACCTCCGGCGAGATCATCAAG 497  
|| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||

RESULT 8  
BE726719

LOCUS BE726719 488 bp mRNA linear EST 14-SEP-2000  
DEFINITION Chlamydomonas reinhardtii CC-1690, normalized, Lambda Zap II  
Chlamydomonas reinhardtii cDNA, mRNA sequence.

ACCESSION BE726719  
VERSION BE726719  
KEYWORDS EST.  
SOURCE BE726719.1 GI:10127903  
ORGANISM Chlamydomonas reinhardtii.

REFERENCE Chlamydomonas reinhardtii  
AUTHORS Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;  
1 (bases 1 to 488) Chlamydomonadaceae; Chlamydomonas.  
Grossman,A., Davies,J., Federspiel,N., Harris,E., Lefebvre,P.,  
McDermott,J.P., Silflow,C., Stern,D. and Surzycki,R.  
Analyses of the Chlamydomonas reinhardtii Genome: A Model,  
TITLE Unicellular System for Analyzing Gene Function and Regulation in  
Vascular Plants; project phase 2  
JOURNAL Unpublished (2000)

Contact: Charles Hauser  
 DCMB Box 91000  
 Duke University  
 Durham, NC 27708-1000  
 Tel: 919 613 8159  
 Fax: 919 613 8177  
 Email: chauser@duke.edu.

FEATURES  
 source  
 Location/Qualifiers  
 1. 488  
 /organism="Chlamydomonas reinhardtii"  
 /strain="CC-1690 wild type mt+ 21gr"  
 /db\_xref="taxon:3055"  
 /clone\_lib="C. reinhardtii CC-1690, normalized, Lambda Zap  
 II"  
 /note="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2:  
 XhoI. This library, constructed by John Davies and Jeffrey  
 McDermott, combines cDNAs from CC-1690 cells grown to  
 mid-log phase in TAP (acetate-containing) medium in the  
 light, TAP medium in the dark, HS (minimal) medium in  
 ambient levels of CO<sub>2</sub> and HS medium bubbled with 5% CO<sub>2</sub>.  
 PolyA mRNA was purified from each sample, pooled and cDNA  
 synthesized. The cDNA was directionally cloned into lambda  
 Zap II (Stratagene) in the EcoRI (5') and XhoI (3') sites.  
 pBluescript II SK- plasmids were excised from the lambda  
 Zap clones by superinfection with EXAssist (Stratagene)  
 phage. The library was normalized using method 4 described  
 in Bonaldo et al (1996) Genome Research 6: 791-806."

BASE COUNT 88 a 177 C 140 G 83 T  
 ORIGIN

Query Match 1.4%; Score 44; DB 12; Length 488;  
 Best Local Similarity 48.1%; Pred. No. 0.37;  
 Matches 125; Conservative 0; Mismatches 135; Indels 0; Gaps 0;

QY 2934 CTTGAGCGATAAACCCGCTCCTGTTTGCACGGCGGGCGTGGACGCCACCTGACGGACG 2993  
 Db 103 CTTGGCGGAGAAGTACACCTCGCTACCCGGCGGCATGTTGCCCGACGTTTCCAGAT 162  
 QY 2994 CGACTACACGGTAACGGCGGCTTTACCGCGCGCACATGCAGCAACCGCAAGACGGGGGC 3053  
 Db 163 CATTGTCAGGAGAAAGGGTGTTCACCAACGGTGACCTGCCCCACCAACGACCAAGCT 222  
 QY 3054 ACGCAATATCCGCACACCCCGCTGTTGCCGGTCTGGCGCGGATGTCGAATTCGGCAA 3113  
 Db 223 CCGCATCTGTGTCAGGTGTGTCGCCCTGGCCCTGCTGATCGAGAAGCGCGGTGCCTC 282  
 QY 3114 CGGCTGGAACGGCTTGGCAGCTTACACCTAGCCGGTTCCAACACGATACGGCAACACAG 3173  
 Db 283 CAGCTGGCGAGCGAAGCGCGTGTCCGCTCTGGACATCCCATCTGTTGTCGACCAACG 342  
 QY 3174 CGGACGAGTGGCGCTAGGCT 3193  
 Db 343 CACCCAGATCTGTCAGGCT 362

RESULT 9  
 LOCUS AV639324  
 DEFINITION AV639324 Chlamydomonas reinhardtii 5% CO<sub>2</sub> Chlamydomonas reinhardtii  
 cDNA clone HC097h07\_r 5', mRNA sequence.  
 ACCESSION AV639324  
 VERSION AV639324.1 GI:10782652  
 KEYWORDS EST.  
 SOURCE Chlamydomonas reinhardtii.  
 ORGANISM Chlamydomonas reinhardtii.  
 Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;  
 Chlamydomonadaceae; Chlamydomonas.  
 REFERENCE 1 (bases 1 to 490)  
 AUTHORS Asamizu E., Miura K., Kucho K., Inoue Y., Fukuzawa H., Ohyama K.,  
 Nakamura Y. and Tabata S.  
 TITLE Generation of expressed sequence tags from low-CO<sub>2</sub> and high-CO<sub>2</sub>  
 adapted cells of Chlamydomonas reinhardtii  
 JOURNAL DNA Res. 7 (5). 305-307 (2000)

```
MEDLINE 20539644
COMMENT Contact: Erika Asamizu
          The First Laboratory for Plant Gene Research
          Kazusa DNA Research Institute
          Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
          Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES
  source      1. .490
              /organism="Chlamydomonas reinhardtii"
              /strain="C9"
              /db_xref="taxon:3055"
              /clone="HC097n07_r"
              /clone_lib="Chlamydomonas reinhardtii 5% CO2"
              /note="vector: pBluescriptII SK-; Site.1: EcoRI; Site.2:
              XhoI; The cDNA library was constructed from cells cultured
              in a medium with bubbling air containing 5% carbon
              dioxide"
BASE COUNT  89 a 176 c 142 g 83 t
ORIGIN
Query Match      1.4%; Score 44; DB 10; Length 490;
Best Local Similarity 48.1%; Pred. No. 0.37;
Matches 125; Conservative 0; Mismatches 135; Indels 0; Gaps 0;

Qy 2934 CTTGAGCGATAAAGCCGTCCTGTTTGAACGGCGGGCGTGGAACGGACCTGAACGGGACG 2993
      ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 98  CTTGGCGGAGAAGTACACCCCTGCGCTACACCGCGGCATGTGCGCGACGTGTTCCAGAT 157
      ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 2994 CGACTACAGGTAAAGCGCGCTTTACCGCGCGGCTGCAACACCGGCAACGCGGGGCG 3053
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 158 CATTGTCTCAAGAGAAGGTTGTTTCACCAACGTGACCTCGCCGCCACCAAGGCCAAGCT 217
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 3054 ACGCAATATGCGGCACACCCCGCTGTTTGAACGGCGGGCGTGGAACGGACCTGAACGGGCA 3113
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 218 CCGCATCTCTGTCGAGGTGCTCCCGCGCCCTGCTGATCGAGAAGCGCGCGGTGCGCTC 277
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 3114 CGGCTGGAACGGCTGGCAGCTTACAGTACGCCCGGTTCCAAACAGTACGGCAACCCACAG 3173
      ||| ||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 278 CAGCTGCGACGGCAAGGCGGTGTCGGCTCTGGACATCCCATCCCTGTTGTCGACCGACGCG 337
      ||| ||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 3174 CGGACGAGTCGGCGTAGGCT 3193
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 338 CACCCAGATCTGCTACGGCT 357
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 10
AV636203 AV636203 Chlamydomonas reinhardtii 5% CO2 Chlamydomonas reinhardtii
LOCUS cDNA clone HC056e01_r 5', mRNA sequence.
DEFINITION
ACCESSION AV636203.1 GI:10779523
VERSION AV636203
KEYWORDS EST.
SOURCE Chlamydomonas reinhardtii.
ORGANISM Chlamydomonas reinhardtii.
          Eukaryote; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
          Chlamydomonadales; Chlamydomonas.
REFERENCE 1 (bases 1 to 497)
AUTHORS Asamizu,E., Miura,K., Kucho,K., Inoue,Y., Fukuzawa,H., Ohyama,K.,
          Nakamura,Y. and Tabata,S.
          Generation of expressed sequence tags from low-CO2 and high-CO2
          adapted cells of Chlamydomonas reinhardtii
          DNA Res. 7 (5), 305-307 (2000)
JOURNAL 20539644
MEDLINE
COMMENT Contact: Erika Asamizu
          The First Laboratory for Plant Gene Research
          Kazusa DNA Research Institute
          Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
          Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES
  source      1. .497
              /organism="Chlamydomonas reinhardtii"
              /strain="C9"
```

```
/db_xref="taxon:3055"
/clone="HC056e01_r"
/clone_lib="Chlamydomonas reinhardtii 5% CO2"
/note="vector: pBluescriptII SK-; Site.1: EcoRI; Site.2:
XhoI; The cDNA library was constructed from cells cultured
in a medium with bubbling air containing 5% carbon
dioxide"
BASE COUNT  90 a 179 c 144 g 84 t
ORIGIN
Query Match      1.4%; Score 44; DB 10; Length 497;
Best Local Similarity 48.1%; Pred. No. 0.37;
Matches 125; Conservative 0; Mismatches 135; Indels 0; Gaps 0;

Qy 2934 CTTGAGCGATAAAGCCGTCCTGTTTGAACGGCGGGCGTGGAACGGACCTGAACGGGACG 2993
      ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 108 CTTGGCGGAGAAGTACACCCCTGCGCTACACCGCGGCATGTGCGCGACGTGTTCCAGAT 167
      ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 2994 CGACTACAGGTAAAGCGCGCTTTACCGCGCGGCTGCAACACCGGCAACGCGGGGCG 3053
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 168 CATTGTCAAGGAGAAGGTTGTTTCACCAACGTGACCTCGCCGCCACCAAGGCCAAGCT 227
      ||| ||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 3054 ACGCAATATGCGGCACACCCCGCTGTTTGAACGGCGGGCGTGGAACGGACCTGAACGGGCA 3113
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 228 CCGCATCTCTGTCGAGGTGCTCCCGCGCCCTGCTGATCGAGAAGCGCGCGGTGCGCTC 287
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 3114 CGGCTGGAACGGCTGGCAGCTTACAGTACGCCCGGTTCCAAACAGTACGGCAACCCACAG 3173
      ||||| ||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 288 CAGCTGCGACGGCAAGGCGGTGTCGGCTCTGGACATCCCATCCCTGTTGTCGACCGACGCG 347
      ||||| ||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 3174 CGGACGAGTCGGCGTAGGCT 3193
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 348 CACCCAGATCTGCTACGGCT 367
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 11
AV636123 AV636123 Chlamydomonas reinhardtii 5% CO2 Chlamydomonas reinhardtii
LOCUS cDNA clone HC055d12_r 5', mRNA sequence.
DEFINITION
ACCESSION AV636123
VERSION AV636123.1 GI:10779443
KEYWORDS EST.
SOURCE Chlamydomonas reinhardtii.
ORGANISM Chlamydomonas reinhardtii.
          Eukaryote; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
          Chlamydomonadales; Chlamydomonas.
REFERENCE 1 (bases 1 to 513)
AUTHORS Asamizu,E., Miura,K., Kucho,K., Inoue,Y., Fukuzawa,H., Ohyama,K.,
          Nakamura,Y. and Tabata,S.
          Generation of expressed sequence tags from low-CO2 and high-CO2
          adapted cells of Chlamydomonas reinhardtii
          DNA Res. 7 (5), 305-307 (2000)
JOURNAL 20539644
MEDLINE
COMMENT Contact: Erika Asamizu
          The First Laboratory for Plant Gene Research
          Kazusa DNA Research Institute
          Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
          Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES
  source      1. .513
              /organism="Chlamydomonas reinhardtii"
              /strain="C9"
              /db_xref="taxon:3055"
              /clone="HC055d12_r"
              /clone_lib="Chlamydomonas reinhardtii 5% CO2"
              /note="vector: pBluescriptII SK-; Site.1: EcoRI; Site.2:
              XhoI; The cDNA library was constructed from cells cultured
              in a medium with bubbling air containing 5% carbon
              dioxide"
BASE COUNT  97 a 181 c 149 g 86 t
ORIGIN
```

Query Match		1.4%;	Score 44;	DB 14;	Length 557;				
Best Local Similarity		48.1%;	Pred. No. 0.39;						
Matches 125;		Conservative	0;	Mismatches 135;	Indels 0;				
Gaps		0;							
QY	2934	CTTTGAGCGATAAAGCGTCTCTGTTTGCACAGCGGGCGGTGGAAACGCGACTGAACGGAGC	2993						
Db	95	CTTGGCGGAGAAGTACACCTCGCGCTACACCGGGCGCATGGTCCCGAGCTGTTCAGAT	154						
QY	2994	CGACTACACGGTAACGGCGGCTTTACCGGGCGGACTGTCAGCAACCGCGCAAGACGGGGC	3053						
Db	155	CATTGTCAGAGGAAGGGTGTGTTCCACCAACGTGACTCGCCACCAACCAAGGCCAAGCT	214						
QY	3054	ACGCAATATGCCGCACACCGCGCTGGTTGCGGGTCTGGGGCGGATGTCGAATTCGGCAA	3113						
Db	215	CGCATCTCTGTTTCGAGGTTGCTCCCTGCGCTGCTGATCGAGAAGCGCGGCGTGCCTC	274						
QY	3114	CGGCTGGAACGGCTTGGCAGCTTACAGCTACGCGGTTCACCAACAGTACGGCAACACAG	3173						
Db	275	CAGCTCGCAGCGCAAGGCGGTGTCGCTCTGGACATCCCCATCTGCTGGTGTGCGACCA	334						
QY	3174	CGGACGAGTCGGCGTAGGCT	3193						
Db	335	CACCCAGATCTGCTACGGCT	354						
RESULT 13									
AV387616									
LOCUS		AV387616	562 bp	mRNA	linear				
DEFINITION		AV387616 Chlamydomonas reinhardtii C9 Chlamydomonas reinhardtii cDNA clone CM027a08_r, mRNA sequence.							
ACCESSION		AV387616							
VERSION		AV387616.1							
KEYWORDS		EST.							
SOURCE		Chlamydomonas reinhardtii.							
ORGANISM		Chlamydomonas reinhardtii.							
REFERENCE		Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales; Chlamydomonadaceae; Chlamydomonas.							
AUTHORS		Asamizu, E., Nakamura, Y., Sato, S., Fukuzawa, H. and Tabata, S.							
TITLE		A large scale structural analysis of cDNAs in a unicellular green alga, Chlamydomonas reinhardtii. I. Generation of 3433 non-redundant expressed sequence tags							
JOURNAL		DNA Res. 6 (6), 369-373 (1999)							
MEDLINE		20152988							
COMMENT		Contact: Yasukazu Nakamura The First Laboratory for Plant Gene Research Kazusa DNA Research Institute Yana 1532-3, Kisarazu, Chiba 292-0812, Japan Email: ynakamu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.							
FEATURES		source							
		1..562							
		/organism="Chlamydomonas reinhardtii"							
		/strain="C9"							
		/db_xref="taxon:3055"							
		/clone_lib="CM027a08_r"							
		/dev_stage="photoautotrophic growth"							
		/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XhoI"							
BASE COUNT		97 a	186 c	173 g	106 t				
ORIGIN									
Query Match									
Best Local Similarity									
Matches 125;									
Conservative									
Mismatches 135;									
Indels									
Gaps									
QY	2934	CTTTGAGCGATAAAGCGTCTCTGTTTGCACAGCGGGCGGTGGAAACGCGACTGAACGGAGC	2993						
Db	95	CTTGGCGGAGAAGTACACCTCGCGCTACACCGGGCGCATGGTCCCGAGCTGTTCAGAT	154						
QY	2994	CGACTACACGGTAACGGCGGCTTTACCGGGCGGACTGTCAGCAACCGCGCAAGACGGGGC	3053						
Db	155	CATTGTCAGAGGAAGGGTGTGTTCCACCAACGTGACTCGCCACCAACCAAGGCCAAGCT	214						
QY	2994	CGACTACACGGTAACGGCGGCTTTACCGGGCGGACTGTCAGCAACCGCGCAAGACGGGGC	3053						
Db	155	CATTGTCAGAGGAAGGGTGTGTTCCACCAACGTGACTCGCCACCAACCAAGGCCAAGCT	214						



```
RESULT 16
CNS017X5
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence Sp6 end of BAC
BAC37LI2 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION
AL108611
VERSION
AL108611.1 GI:5628915
KEYWORDS
GSS.
SOURCE
Drosophila melanogaster.
ORGANISM
Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 1101)
Genoscope.
Direct Submission
Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
Bp 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT
Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelobAC11.
FEATURES
Location/Qualifiers
1..1101
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="BAC37LI2"
/plasmid_lib="DrosBAC"
/plasmid="pBelobAC11"
/note="end : SP6"
BASE COUNT 273 a 245 c 164 g 94 t 325 others
ORIGIN
Query Match 1.4%; Score 43.8; DB 17; Length 1101;
Best Local Similarity 22.5%; Pred. No. 0.61;
Matches 71; Conservative 115; Mismatches 128; Indels 1; Gaps 1;
QY 2048 CACGGAGCTTCGGCAGCGGCACATTCGCCGCCCGCTCTGAACACGCCGTAGAAC 2107
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 779 CAACAARMSNAGCAVCMARVMAVMAVMAGVYMMCAVCAVCMCGGCAMCACSSGGGGRGV 838
QY 2108 AGGCGCGCAGCAATCTGGAACCTGATGCTCGAATCGATGCTCGGAATCATCCGCAA 2167
:|:| | | | | : : : : : : : : : : : : : : : : : : : : : : : :
Db 839 MRGGGGGGMAAASASARAMRAAASAGRCRCMMAARGCRCACARGGVRCRCAMCAM 898
QY 2168 CACCGGAGAC-GGTTGAACTGCGCGCGCGCGACGACAGATATGCGGCGCATCCGCC 2226
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 899 MMAGCSRACAAVMSAGCMRRSRMSGSCMMCMMAACARVRCACARRGSRVCMCAACRC 958
QY 2227 TAGCGGCGCACTTTCGCGCAGCGCGCGGTACACATCGCGATGCCGCGCAGGTGTA 2286
:|:| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 959 MVRCSRVSVSACMCRCASGAGCASSAGMSRCMMCMRRSRGSGGCGACVCAVMA 1018
QY 2287 CGCATCTTCAACAGTCTCGCGCTACCGTCTATGCGCAGACAGTAGTCCGCGCGCCATCGCGAT 2346
:|:| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1019 VVCARMMAVAVMSGRSRSCHCAMMACAMCAMVAMRCVAVAGCAASCMVMVMACAC 1078
QY 2347 ATCAGGGAGCGCGG 2361
: | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1079 MMGCASRCACASVRS 1093
RESULT 17
AW676029
LOCUS
DEFINITION
832001E09.Y1 C. reinhardtii CC-125 nutrient replete, -S, -Fe,
Lambda Zap II Chlamydomonas reinhardtii cDNA similar to Sbp2 gene
encoding seduheptulose-1,7-bisphosphatase, mRNA sequence.
ACCESSION
AW676029
VERSION
AW676029.1 GI:7545541
KEYWORDS
EST.
SOURCE
Chlamydomonas reinhardtii.
ORGANISM
Chlamydomonas reinhardtii.
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.
REFERENCE
1 (bases 1 to 655)
Grossman,A., Davies,J., Federspiel,N., Harris,E., Lefebvre,P.,
Silflow,C., Stern,D. and Surzycki,R.
Analyses of the Chlamydomonas reinhardtii Genome: A Model.
Unicellular System for Analyzing Gene Function and Regulation in
Vascular Plants
Unpublished (2000)
Contact: Elizabeth H. Harris
DCMB Box 91000
Durham, NC 27708-1000, USA
Tel: 919 613 8164
Fax: 919 613 8177
Email: chlamy@duke.edu
Sbp2 (Csbp) gene; compare C. reinhardtii Accession X74418.
FEATURES
Location/Qualifiers
1..655
/organism="Chlamydomonas reinhardtii"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-125 nutrient replete, -S,
-Fe, Lambda Zap II"
/note="Vector: pBluescript II SK-; Site_1: EcoRI; Site_2:
XhoI; This library combines cDNAs from nutrient replete
(TAP medium, mid-log phase), sulfur-deprived (TAP medium
-S for 2 hr), and iron-deprived cells from CC-125.
Fe-deprived cells were obtained by either (1) growth in
SGII medium at 0.18 micromolar Fe, 2X wash with SGII -Fe
and resuspension in SGII -Fe + 80 micromolar
bathophenanthroline sulfonic acid (an iron chelator) for 2
hours or (2) growth in SGII medium, 2X wash with SGII -Fe,
resuspension in SGII -Fe and continued growth for 48
hours. 1.25 micrograms of each polyA mRNA prep was used to
synthesize cDNA which was directionally cloned into lambda
ZAP II (Stratagene) in the EcoRI (5') and XhoI (3') sites.
pBluescript II SK- plasmids were excised from the lambda
ZAP clones by superinfection with EXassist (Stratagene)
phage."
BASE COUNT 117 a 234 c 192 g 112 t
ORIGIN
Query Match 1.3%; Score 43.2; DB 10; Length 655;
Best Local Similarity 49.6%; Pred. No. 0.72;
Matches 111; Conservative 0; Mismatches 113; Indels 0; Gaps 0;
QY 2934 CTTGAGCGATAAAGCGTCTCTGTTGCAACGCGGCGGTGGAACGCGACTGAACGGACG 2993
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 289 CTTGGCGGAGAACTACACCTCGCTACACCGCGGATGTCGCCGACGTGTCAGAT 348
QY 2994 CGACTACAGGTAAACGGCGGCTTTACCGCGCGACTGTCAGCAACCGCGAAGACGGGGG 3053
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 349 CATTTGTCAAGGAGAGGGTGTGTTTCAACACGTGACCTCGCCACCACCAAGCCAAAGCT 408
QY 3054 ACCCAATATGCCGACACCGCGCTGTTGCGGTCTGGGCGGGATGTGCAATTCGGCAA 3113
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 409 CCGCATCTCTTTCGAGGTTGCTCCCTGCGCTGTCGATCGAAGGCCGCGGTGCTC 468
QY 3114 CGGCTGGAACGGCTTGCACGTTACAGCTACGCGGTTCACAAAC 3157
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 469 CAGCTGGCGGCAAGGCGGTGTCGCTCTGGACATCCCCATC 512
RESULT 18
```

AV633578  
LOCUS AV633578 477 bp mRNA linear EST 15-DEC-2000  
DEFINITION AV633578 Chlamydomonas reinhardtii 5% CO<sub>2</sub> Chlamydomonas reinhardtii  
cDNA clone HC022e07\_r 5', mRNA sequence.

ACCESSION AV633578  
VERSION AV633578.1 GI:10776898

KEYWORDS EST.  
SOURCE Chlamydomonas reinhardtii.

ORGANISM Chlamydomonas reinhardtii  
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;  
Chlamydomonadaceae; Chlamydomonas.

REFERENCE 1 (bases 1 to 477)

AUTHORS Asamizu, E., Miura, K., Kucho, K., Inoue, Y., Fukuzawa, H., Ohyama, K., Nakamura, Y. and Tabata, S.

TITLE Generation of expressed sequence tags from low-CO<sub>2</sub> and high-CO<sub>2</sub>

JOURNAL adapted cells of Chlamydomonas reinhardtii

MEDLINE DNA Res. 7 (5), 305-307 (2000)

COMMENT 20539644

Contact: Erika Asamizu

The First Laboratory for Plant Gene Research

Kazusa DNA Research Institute

Yana 1532-3, Kisarazu, Chiba 292-0812, Japan

Email: asamizu@kazusa.or.jp, URL: <http://www.kazusa.or.jp/en/plant/>.

FEATURES  
Location/Qualifiers

1..477

/organism="Chlamydomonas reinhardtii"

/strain="C9"

/db\_xref="taxon:3055"

/clone="HC022e07\_r"

/clone\_lib="Chlamydomonas reinhardtii 5% CO<sub>2</sub>"

/note="Vector: pBluescriptII SK-; Site\_1: EcoRI; Site\_2: XhoI; The cDNA library was constructed from cells cultured in a medium with bubbling air containing 5% carbon dioxide"

BASE COUNT 87 a 168 c 138 g 83 t 1 others

ORIGIN

Query Match 1.3%; Score 43; DB 10; Length 477;

Best Local Similarity 47.7%; Pred. No. 0.71;

Matches 124; Conservative 0; Mismatches 136; Indels 0; Gaps 0;

QY 2934 CTTGACGATTAAGCGTCTGTTTGCAACGCGGGCGGTGGACGGAGCTCAACGGAGC 2993

DB 211 CTGGCGGAGAGTACACCTCGCTACACCGCGGCGATGTGCGCGAGCTGTCCAGAT 270

QY 2994 CGACTACAGGTAAACGGGGCGCTTTACCGCGCGGCTGACGACCAACGGCAACGGGGC 3053

DB 271 CATTTGACAGGAGAGGGTGTGTTACCAACGTGACCTGCCCCACCACCAAGCCAAAGCT 330

QY 3054 ACGCAATATGCCCGACACCCCGCTGTTGCCGGTCTGGGCGCGGATGTGGAATTCGGCAA 3113

DB 331 CGGCATCTCTTCGAGGTGCTCCCTGCGCTGATCGAGAGCGCGGGTGCCTC 390

QY 3114 CGCTGGACGCGCTGGCAGTTACAGCTACCCCGGTTCCAAACAGTACGGCAACACAG 3173

DB 391 CAGCTGCGACGCAAGGCGGTGCTGCGCTCTGGACATCCCATCTGGTGTGGCAGCAG 450

QY 3174 CGGACGAGTCGCGTAGGCT 3193

DB 451 CACCCAGATCTCTACGGCT 470

RESULT 19

LOCUS BG947100

DEFINITION IP1\_2\_E12.bl\_A002 Immature pannicle 1 (IP1) Sorghum bicolor CDNA,

mRNA sequence.

ACCESSION BG947100

VERSION BG947100.1 GI:14365289

KEYWORDS EST.

SOURCE sorghum.

ORGANISM Sorghum bicolor

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
clade; Panicoideae; Andropogoneae; Sorghum.

1 (bases 1 to 358)

AUTHORS Klein, R.R., Cordonnier-Pratt, M.-M., Gingle, A., Sudman, M. and Pratt

L.H.

TITLE An EST database from Sorghum: developing preanthesis pannicles

JOURNAL Unpublished (2001)

COMMENT Contact: Cordonnier-Pratt MM

Laboratory for Genomics and Bioinformatics

The University of Georgia, Department of Plant Biology

Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA

Tel: 706 542 1860

Fax: 706 583 0210

Email: mmpratt@uga.edu

Sequences have been trimmed to exclude polyA, vector and regions  
below Phred quality 16. The threshold for high quality sequence is

20. Three-prime sequences, which are obtained with PolyMix or T7

sequencing primer, are presented as the reverse complement.

Seq primer: JEN REV

High quality sequence stop: 352

POLYA-No.

FEATURES

source

1..358

/organism="Sorghum bicolor"

/cultivar="BTx623"

/db\_xref="taxon:4558"

/clone\_lib="Immature pannicle 1 (IP1)"

/note="Organ: Developing preanthesis pannicles; Vector:  
pBluescript II SK(-) from Lambda Zap II; Site\_1: XhoI;  
Site\_2: EcoRI; The library was made from poly-A RNA in the  
cloning vector lambda Zap II. Clones to be sequenced were  
prepared by mass excision."

BASE COUNT 55 a 129 c 128 g 46 t

ORIGIN

Query Match 1.3%; Score 42.2; DB 13; Length 358;

Best Local Similarity 55.0%; Pred. No. 1;

Matches 83; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 1219 ATTGCGGAACATCTTTTCGACCCATCGTAACCGGCACGCGGCTGCTGCTGTCGAG 1278

DB 24 ATCTCGGAGACGTCCATGCTGTCGCCGACGTCGCGGCTGCTCCGCGG 83

QY 1279 AAATACCGGTGATGAGCAACGACAACTGCGTACCACTGCTGACAAAGGCTCAGGAC 1338

DB 84 GGCACCCCGAGTGGAGCCCGCAGCCATCCGCTCGGCTGATGACGACGCGTACAAC 143

QY 1339 ATCGGTGACGTGCGGCTGGACAGCAAGTTTCG 1369

DB 144 GATTACCCAGCGCGCTGGCATCTCTGGACG 174

RESULT 20

BE041076/c

LOCUS

DEFINITION

OF19D09 OF Oryza sativa cDNA 5' similar to glycine-rich protein,

mRNA sequence.

ACCESSION

BE041076

VERSION

BE041076.1

KEYWORDS

EST.

SOURCE

Oryza sativa.

ORGANISM

Oryza sativa

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Ehrhartoideae; Oryzeae; Oryza.

REFERENCE

AUTHORS

1 (bases 1 to 755)

Bohnert, H.J., Borchert, C., Brazille, S., Brooks, J., Eaton, M., Ferrea

, H., Kawasaki, S., McCallough, A., Michalowski, C.B., Palacio, C.,

Scara, G., Wheeler, M. and Zepeda, G.R.

Functional Genomics of Plant Stress Tolerance

Unpublished (2000)

Contact: Michalowski, C.B.

University of Arizona

Bio Sciences West room 513, Tucson, AZ 85721, USA  
Tel: 520-621-7982  
Fax: 520-621-1697  
Email: cbm@u.arizona.edu  
An open reading frame exists.

**FEATURES**  
**source**

```

source
1 735
/organism="Oryza sativa"
/strain="pokali"
/db_xref="taxon:4530"
/clone_lib="of"
/tissue_type="entire plant"
/dev_stage="2 weeks"
/note="1 week 150mM NaCl"
105 a 216 c 258 g 176 t
BASE COUNT

```

Query Match	1.38;	Score	42.2;	DB	10;	Length	755;
Best Local Similarity	47.5%;	Pred.	No. 1.5;				
Matches	125;	Conservative	0;	Mismatches	138;	Indels	0; Gaps
Qy	2109	GGGCGGCAGCAATCTGGAAAAACCTGATGGTCGAACATGGATCCGCCAATCATCGCACAAC	2168				
Dd	515	GGCCCCACCACATCAGGTCTCCAGGTCGCCGCGGAGTCCGCCACGGAGCCACCACCCAC	456				
Qy	2169	ACCCGAGCAGGTTGAAACTGGGGCGCGGACCGCACAGATATGCGGGGCATCCGGCCCCCTA	2228				
Dd	455	GACTGCCACCGTAGCCGCGCGCGCCGCGCCAGCCGCGCAGTAGCGCGCGCGCGCCAC	396				
Qy	2229	C GGCGCAACTTTC CGCGCAGCGGCAGCCGTACAGCATCGGAATCGCGCGCAGCGGTGTACG	2288				
Dd	395	CGTAGCGCCCTCAGGGCGCCGTAGCGCGCGCGCCACCTCTCCGTAGCGCGCGCCTC	336				
Qy	2289	CATCTTGCAACAGTCTGCCCGTACCGTCTATGCGGACAGTAGCCGCGCGCCATGCGGATAT	2348				
Dd	335	CACGACCGCCCGGTAGCGGCCACCGCGCGCGCGCGGTAGCCCCCGGCTCCGCGCGGGAGC	276				
Qy	2349	GCAGGGACGCCGCTGAAAGCCG	2371				
Dd	275	GGCGGGACTGGGCGCTCATTGACG	253				

RESULT 21

CNS017GS/c	CNS017GS	1101 bp	DNA	linear	GSS 26-JUL-1999
LOCUS	Drosophila melanogaster genome survey sequence SP6 end of BAC				
DEFINITION	BACN17J04 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.				
ACCESSION	AL108022				
VERSION	AL108022.1				
KEYWORDS	GI:5628326				
SOURCE	GSS.				
ORGANISM	Drosophila melanogaster.				
	Drosophila melanogaster				
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;				
	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;				
REFERENCE	Ephyraoidea; Drosophilidae; Drosophila.				
	1 (bases 1 to 1101)				
	Genoscope.				
	Direct Submission				
AUTHORS	Submitted (23-JUL-1999)				
TITLE	Genoscope - Centre National de Sequencage :				
JOURNAL	Submitted (23-JUL-1999)				

COMMENT

near : [www.genoscope.cns.fr/17](http://www.genoscope.cns.fr/17)

Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - <http://www.edgp.ebi.ac.uk> -. This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelOBAC11.

```

FEATURES
source
Location/Qualifiers
1. .1101
/orqanism="Drosophila melanogaster"

```

```
/db_xref="taxon:7227"  
/clone="BACN17J04"  
/clone_lib="DrosBAC"  
/plasmid="pBelobAC11"  
/note="end: SP6"
```

BASE COUNT	268 a	126 c	100 g	195 t	412 others
ORIGIN					

BASE COUNT	258 a	126 c	100 g	195 t	412 Others
Query Match	1.3%	Score 42;	DB 17;	Length 1101;	
Best Local Similarity	19.3%;	Pred. No. 2;			
Matches	80;	Conservative 148;	Mismatches 187;	Indels 0;	Gaps 0;
Qy	254	CGGTACAGACAGGATCCCAAAATCAATGCCCCCCCCCGGAATCTGCATACCGGAGACT	313		
Db	1087	CVSSYMGAGACGCMRRYVMSRRMAAVAVAMAGMMAMAAAAAAMVRRGAGVCGRVMGVSS	1028		
Qy	314	TTACAAACCCAAATGACGCATACAGAAATTTGATCAACCTCAACTTCAATTTGAAGCAG	373		
Db	1027	VGMSGSCSSMRVVCRCMCMAMMMVSMRAAAAVRMCAAAAAMMAMAMAAVARM	968		
Qy	374	GCTATACAGGACGGGGTAGAGTAGTATCGTCGTATACAGGCGAAATCCGTCGGCAGCA	433		
Db	967	AAAAMASGAARGGRVRRGAMAGSSVSCGCGVYVRCBVRVMSAMMMVMVMCMCSMMAA	908		
Qy	434	TATCCTTTCCCGAACTGTATGCAGAAAGAACAGCGGTATACAGCAAAATTACAAAAACT	493		
Db	907	AMAMMAAMVMVMSAAATYVHARAGRRVRRRRDRVSSVYVRRGRGAAAAAGA	848		
Qy	494	ATACGGCGTATATCGGGAAGGAGCGCTCGAAGACGGAGCGGTAAAGACATTTAAAGCTT	553		
Db	847	ACMYVCCCVVYGRAAAAAARAAACVSCVCMCMCAAAMVAMCMCAAAARSGRSGARMSS	788		
Qy	554	CTTTCCGACGTATGAGCCGCTTATAGACACTGAAGCAAGCCGAGGATATCCGCCACGTAA	613		
Db	787	ASHTTTTDBRAGSSACGAAAACRVARSVGGACCAACSMGRSRRGAGWTTTKDDWDATAR	728		
Qy	614	AAGAAATCGGACACATCATGATCGTGTCTCCCATATATTATTTGGCGGGCGTTCCGTTGA	668		
Db	727	RAMAAATKKNNWRATTTTWTGTTTAAACAMWTRKTTTITTTTMAASCVTVTRTADA	673		

RESULT 22

CNS017SY	CNS017SY	1101 bp	DNA	linear	GSS 26-JUL-1999
LOCUS					
DEFINITION	Drosophila melanogaster genome survey sequence Sf6 end of BAC BAC37L08 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.				
ACCESSION	AL108460				
VERSION	AL108460.1	GI:5628764			
KEYWORDS	GSS.				
SOURCE	Drosophila melanogaster.				
ORGANISM	Drosophila melanogaster.				
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				
REFERENCE	1 (bases 1 to 1101)				

**TITLE**

JOURNAL	COMMENT
Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - FRANCE (E-mail : <a href="mailto:secref@genoscope.cns.fr">secref@genoscope.cns.fr</a> - Web : <a href="http://www.genoscope.cns.fr">www.genoscope.cns.fr</a> )	Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - <a href="http://www.edgp.ebi.ac.uk">http://www.edgp.ebi.ac.uk</a> . This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBeloBAC11.

FEATURES  
SOURCE

FEATURES	Location/Qualifiers
source	1. .1101
	/organism="Drosophila melanogaster" /db xref="taxon:7227"





to give pBluescript phagemids before subtraction was carried out. The mass excision of phagemid library and subtraction were done in HT Nguyen lab by D. Zhang at Texas Tech University. Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."

BASE COUNT 109 a 198 c 145 g 86 t  
ORIGIN

Query Match 1.3%; Score 41.8; DB 14; Length 538;  
Best Local Similarity 54.1%; Pred. No. 1.7; Indels 0; Gaps 0;

Matches 85; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 2941 GATAAAGCGCTGCTGTTTTCACGCGCGGTGGAAACGGACCTGTAACGGACGCGACTAC 3000

Db 397 GTTGAAGCGCGCGGTGATCGCGGACGTGGAGGAGACCTACTCGGCACCGGAGAC 338

QY 3001 ACGGTAACGGCGGCGTTTACCGCGCGGACTGCAGCAACCGGAAGACGGGGCGACGCAAT 3060

Db 337 CTGGTCCATGGCGGCGTGGCGCGGACATGAGCTTGGTGGCAGACCGAGCTTGGCGGTG 278

QY 3061 ATGCCGCACACCGCTGCTGCCGTCTGGGCGCG 3097

Db 277 AGACGGAGACCGCCAGGAGGTGATGTGGCGGTG 241

RESULT 25

AL721811

LOCUS AL721811 650 bp mRNA linear EST 18-APR-2002  
DEFINITION AL721811 Danio rerio embryonic inner ear subtracted cDNA Danio  
ACCESSION AL721811  
VERSION AL721811.1 GI:20186415  
KEYWORDS EST  
SOURCE zebrafish.

ORGANISM

Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes  
; Cyprinidae; Danio.  
REFERENCE 1 (bases 1 to 650)  
AUTHORS Coimbra,R., Weil,D., Brottier,P., Blanchard,S., Levi,M., Hardelin  
J.P., Weissenbach,J. and Petit,C.  
TITLE A subtracted cDNA library from the zebrafish (Danio rerio)  
JOURNAL embryonic inner ear  
COMMENT Unpublished (2002)  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES

source

1..650  
/organism="Danio rerio"  
/db\_xref="taxon:7955"  
/clone="BN0AA0542A05"  
/clone\_lib="Danio rerio embryonic inner ear subtracted  
cDNA"  
/tissue\_type="inner ear"  
/dev\_stage="embryonic"  
/note="subtracted cDNA library"

BASE COUNT 213 a 122 c 160 g 155 t  
ORIGIN

Query Match 1.3%; Score 41.8; DB 9; Length 650;  
Best Local Similarity 51.3%; Pred. No. 1.8; Indels 0; Gaps 0;

Matches 97; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 463 GAACACGGCTATACGAAATTTACAAAACTATACGGCGTATATCGGGAAGGAGCGCCT 522

Db 331 GTAATGGTGAAGAAAGTGAAGTTCATCTAGTTTCAGAAATGCAGAAATGCATCATCT 390

QY 523 GAAGACGGAGCGGTAAAGACATTAAGCTCTTTTCGACGATGAGCGCCCTTATAGAGACT 582

Db 391 AAACAGGTTGTAGTGAACACAAAGAAGCTTTTGATCCAAAGAAACAGCATTTGAAGAGGAT 450

QY 583 GAAGCAAAGCCGACGGATATCCGCCACGTAAAGAAATCGACACATCGATGTGGTCTCC 642  
Db 451 GAACAATGGATAAAGAGAGAGTCCATTAAAGGATCGCTCCATAAGACTTCCAGC 510  
QY 643 CATATTATT 651  
Db 511 CCTATTGTT 519

RESULT 26

AL721550

LOCUS AL721550 766 bp mRNA linear EST 18-APR-2002  
DEFINITION AL721550 Danio rerio embryonic inner ear subtracted cDNA Danio  
ACCESSION AL721550  
VERSION AL721550.1 GI:20186154  
KEYWORDS EST.  
SOURCE zebrafish.

ORGANISM

Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes  
; Cyprinidae; Danio.  
REFERENCE 1 (bases 1 to 766)  
AUTHORS Coimbra,R., Weil,D., Brottier,P., Blanchard,S., Levi,M., Hardelin  
J.P., Weissenbach,J. and Petit,C.  
TITLE A subtracted cDNA library from the zebrafish (Danio rerio)  
JOURNAL embryonic inner ear  
COMMENT Unpublished (2002)  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES

source

1..766  
/organism="Danio rerio"  
/db\_xref="taxon:7955"  
/clone="BN0AA0532A12"  
/clone\_lib="Danio rerio embryonic inner ear subtracted  
cDNA"  
/tissue\_type="inner ear"  
/dev\_stage="embryonic"  
/note="subtracted cDNA library"

BASE COUNT 243 a 147 c 194 g 182 t  
ORIGIN

Query Match 1.3%; Score 41.8; DB 9; Length 766;  
Best Local Similarity 51.3%; Pred. No. 2; Indels 0; Gaps 0;

Matches 97; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 463 GAACACGGCTATACGAAATTTACAAAACTATACGGCGTATATCGGGAAGGAGCGCCT 522

Db 376 GTAATGGTGAAGAAAGTGAAGTTCATCTAGTTTCAGAAATGCAGAAATGCATCATCT 435

QY 523 GAAGACGGAGCGGTAAAGACATTAAGCTCTTTTCGACGATGAGCGCGTATAGAGACT 582

Db 436 AAACAGGTTGTAGTGAACACAAAGAAGCTTTTGATCCAAAGAAACAGCATTTGAAGAGGAT 495

QY 583 GAAGCAAAGCCGACGGATATCCGCCACGTAAAGAAATCGACACATCGATGTGGTCTCC 642

Db 496 GAACAATGGATAAAGAGAGAGTCCATTAAAGGATCGCTCCATAAGACTTCCAGC 555

QY 643 CATATTATT 651

Db 556 CCTATTGTT 564

RESULT 27

AL716995

LOCUS AL716995 772 bp mRNA linear EST 18-APR-2002  
DEFINITION AL716995 Danio rerio embryonic inner ear subtracted cDNA Danio  
ACCESSION AL716995  
rerio cDNA clone BN0AA0212C10 3', mRNA sequence.

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES  
source

Location/Qualifiers  
1. .790  
/organism="Danio rerio"  
/db\_xref="taxon:7955"  
/clone="BN0AA026ZG05"  
/clone\_lib="Danio rerio embryonic inner ear subtracted cDNA"  
/tissue\_type="inner ear"  
/dev\_stage="embryonic"  
/note="subtracted cDNA library"

BASE COUNT      264 a    153 c    192 g    181 t

ORIGIN

Query Match                  1.3%; Score 41.8; DB 9; Length 790;  
Best Local Similarity       51.3%; Pred. No. 2;  
Matches     97; Conservative    0; Mismatches    92; Indels     0; Gaps     0;

Qy 463 GAACAGCGCTATACGAAAATTACAAAACATTATAGGGGTATATGCGGAAGAACGCCCT 522  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 333 GTAAATGCTGAAAAAGAAGTGAAGTTTCATCTAGTTCCAGAAATGCAGATGCATCATCT 392  
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Qy 523 GAAGACGAGCGCGTAAAGACATTAAAGCTTCTTCCAGCAGATGAGCGGTATAGAGACT 582  
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 393 AAACAGGTTGTAGTGGAACAAGAAGCTTTGATCCCAAGAAAACAGCAATTGAAGAGGAT 452  
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Qy 583 GAAGCAAAGCGCGGATATATCCGCCACGTAAAGAAATCGGACACATCGATGTGCTCTCC 642  
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 453 GAACAAATGGATAAAGAGAGAGTCCATTTAAAGGAATCGTCCCATAAAGACTTCCAGC 512  
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Qy 643 CATATTATT 651  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 513 CCTATTGTT 521  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 29

AL717085/c

LOCUS

DEFINITION AL717085 Danio rerio embryonic inner ear subtracted cDNA Danio  
rerio cDNA clone BN0AA021ZC10 5', mRNA sequence.

ACCESSION AL717085

VERSION AL717085.1 GI:20181688

KEYWORDS EST.

SOURCE zebrafish.

ORGANISM Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes  
; Cyprinidae; Danio.  
1 (bases 1 to 792)  
Colmбра,R., Weil,D., Brottier,P., Blanchard,S., Levi,M., Hardelin,  
J.P., Wesslenbach,J. and Petit,C.  
A subtracted cDNA library from the zebrafish (Danio rerio)  
embryonic inner ear  
Unpublished (2002)  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES  
source

Location/Qualifiers  
1. .792  
/organism="Danio rerio"  
/db\_xref="taxon:7955"  
/clone="BN0AA021ZC10"  
/clone\_lib="Danio rerio embryonic inner ear subtracted cDNA"  
/tissue\_type="inner ear"  
/dev\_stage="embryonic"  
/note="subtracted cDNA library"

BASE COUNT      177 a    194 c    165 g    256 t

ORIGIN

Query Match                  1.3%; Score 41.8; DB 9; Length 792;  
Best Local Similarity       51.3%; Pred.No. 2;



```
Db 525 NNSNNNSNSSATNNNNCTANAANASSANNNTSAANGSSNNSSNNAASASNTT 466
QY 2032 CGTCGGCGCAATCGGCACCGAGTCTTCGGCAGCGGCACATTCCGGCGCGCGGTCG 2091
Db 465 NSSSNCNNNNNSNSTSSNAAAASSSSCASNSSSSASASNAASSSSASANSNSN 406
QY 2092 AAACACCCCTAGAACAGCGCGGAGCAATCTTGAAAACCTGATGTCGAACCTGGATGCC 2151
Db 405 SASSASANSASSSSSSSSSCSNASSAATSNNSNSNSAVSNNSNNNNNNSSSSSS 346
QY 2152 TCCGAATATCCGCACACCGGAGACGGTTGAACCTCGGCGCGCCGACCGACAGATATG 2211
Db 345 SSSSCVAACSSSSSSSSCATSNNSASNSANAASASSSSNNNSNNNSNNNN 286
QY 2212 CCGGGCATCCGCCCTAGCGGCACTTTCCGCGCAGCGGCGACGCTACAGCATGCCAAT 2271
Db 285 SNANNNSNSANSSSSSSCSCSNNSNSNSSSCCSNNSNSSSSSNNNNNSNNNSN 226
QY 2272 GCCCGCAGCGGTACGATCTTCAACAGTCTCGCCGCTATGTCGCGACAGTACC 2331
Db 225 NNNANNNSSSSSSNAANANNNGNAGSGGGSNNNNCCNNAACCNANNAANNNC 166
QY 2332 GCCGCCCATGCC 2343
Db 165 CNNNNCCCCC 154

RESULT 32
LOCUS B0759294
DEFINITION B0759294 444 bp mRNA linear EST 26-JUL-2002.
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
; Triticeae; Hordeum.
B0759294
B0759294.1 GI:21967766
EST.
SOURCE Hordeum vulgare.
ORGANISM Hordeum vulgare
REFERENCE 1 (bases 1 to 444)
AUTHORS Hedley, P., Liu, H., Caldwell, D., McCallum, N., Mudie, S., Cardle, L.,
TITLE Ramsay, L., Machray, G., Marshall, D.F.M. and Waugh, R.
JOURNAL Development of Barley Transcriptome Resources
COMMENT Unpublished (2001)
Contact: Waugh R, Marshall DF
Genome Dynamics/Computational Biology
Scottish Crop Research Institute
Invergowrie, Dundee, DD2 5DA, Scotland, UK
Tel: 00 44 1382 562731
Fax: 00 44 1382 562426
Email: est@scri.sari.ac.uk.
FEATURES
Location/Qualifiers
source
1..444
/organism="Hordeum vulgare"
/db_xref="taxon:4513"
/clone="EBp01_S0004_H04"
/tissue_type="pistil"
/dev_stage="1 DPA"
/lab_host="DH10B"
/note="vector: pSPORT1; site_1: Sal I; site_2: Not I;
Non-normalised library, directionally cloned into pSPORT1.
Derived from pistils dissected from developing grains (24
hours post anthesis) in glasshouse grown barley plants.
Developed as part of the barley transcriptome resources of
BBSRC/SEERAD funded cereal IGF (Investigating Gene
Function) project."
BASE COUNT 74 a 177 c 137 g 56 t
ORIGIN

Query Match 1.3%; Score 41.4; DB 14; Length 444;
```

```
Best Local Similarity 55.1%; Pred. No. 2; Mismatches 0; Gaps 0;
Matches 81; Conservative 0; Indels 66; Indels 0; Gaps 0;

QY 1219 ATTCCGCGAATCTTTTCCGACCATCGTTAAACCGGACGGGCTCTGCTGCTCAG 1278
Db 235 ATTCGCGGACGTCCATGTCGTCGCCACGTGAGCGGCTCGCCGGCTGCTCCGCGAG 294
QY 1279 AAATACCCGTGGATGAGCAACAGACAACTCGGTACCACTGCTGACACAGCGGTCTAGAC 1338
Db 295 GCGACCCGACTGGAGCCCGCGGAGGTCAAGTCGCGCTCATGACCAACCGCGGTACAAC 354
QY 1339 ATCGGTGCATCTCGCGTGGAGACGAAG 1365
Db 355 CTGACAACTCCGGGGAGATCATCAAG 381

RESULT 33
LOCUS BE778605
DEFINITION BE778605 976 bp mRNA linear EST 20-OCT-2000
601466144F1 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:3869242 5',
mRNA sequence.
ACCESSION BE778605
VERSION BE778605.1 GI:10199803
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 976)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9618 row: h column: 11
High quality sequence start: 7
High quality sequence stop: 561.
FEATURES
Location/Qualifiers
source
1..976
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3869242"
/clone_lib="NIH_MGC_67"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Library constructed by Life
Technologies."
BASE COUNT 259 a 331 c 276 g 110 t
ORIGIN

Query Match 1.3%; Score 41.4; DB 12; Length 976;
Best Local Similarity 50.8%; Pred. No. 2.9; Mismatches 0; Gaps 0;
Matches 99; Conservative 0; Indels 96; Indels 0; Gaps 0;

QY 2375 CGGACGGGTGTGACCACAAACGCTACGGGTCTGCGGCTCATCGCGCAACCCCAACAGGACG 2434
Db 730 CGGACCGGTAGCGGGTCTGCCACAGCGCTACGCCCGCGGCTCAGGACCCAGCGG 789
QY 2435 GTGCAACCTGGGAACAGCGGCTGTGTAAGGCAAAATCGCGGGAGTACCCAAACCGTCG 2494
Db 790 ACGGACCGGTGTGAACATAGCAGCGCGCCGCCAACCAACACAGGACGAACACAGCGCGCAC 849
QY 2495 GCATTGCCCGGAAACCGCGGGAATACGACAGCAGCGCGCCACACTGGGACACACA 2554
Db 2554 GCATTGCCCGGAAACCGCGGGAATACGACAGCAGCGCGCCACACTGGGACACACA 2554

Query Match 1.3%; Score 41.4; DB 14; Length 444;
```

```

Db 850 CCAAGCGACACACACACACACACCAACCAACAAACCAACAAACGCGCGAAACACA 909
QY 2555 GCACATGGAGCGAAA 2569
||||| ||| |||
Db 910 ACACACAGCGCCCAA 924

RESULT 34
LOCUS BI778455
DEFINITION EBr007_SQ003_F16_R root, 3 week, reduced light, cv Optic, EBr007
HORDEUM VULGARE cDNA clone EBr007_SQ003_F16 5', mRNA sequence.
ACCESSION BI778455
VERSION 1
KEYWORDS EST.
SOURCE Hordeum vulgare.
ORGANISM Hordeum vulgare.

REFERENCE 1 (bases 1 to 588)
AUTHORS Hedley,P., Liu,H., Caldwell,D., McCallum,N., Mudie,S., Cardle,L.,
Ramsay,L., Machray,G., Marshall,D.F.M. and Waugh,R.
TITLE Development of Barley Transcriptome Resources
JOURNAL Unpublished (2001)
COMMENT On Sep 26, 2001 this sequence version replaced gi:15781347.
Contact: Waugh R. Marshall DF
Genome Dynamics/Computational Biology
Scottish Crop Research Institute
Invergowrie, Dundee, DD2 5DA, Scotland, UK
Tel: 00 44 1382 562731
Fax: 00 44 1382 562426
Email: est@scri.sari.ac.uk
All sequence has a phred quality score of 20 or over
Seq primer: M13 reverse.

FEATURES
source
1..588
Location/Qualifiers
/organism="Hordeum vulgare"
/cultivar="Optic"
/db_xref="taxon:4513"
/clone_lib="root, 3 week, reduced light, cv Optic, EBr007"
/tissue_type="root"
/dev_stage="3 week"
/lab_host="DH10B"
/notes="Vector: pSPORT1; Site_1: Sal I; Site_2: Not I;
Non-normalised library, directionally cloned into pSPORT1.
Derived from shoot and root material of 3 week old
etiolated barley plants. Developed as part of the barley
transcriptome resources of BBSRC/SERAD funded cereal IGF
(Investigating Gene Function) project."

BASE COUNT 102 a 205 c 199 g 82 t
ORIGIN

Query Match 1.3%; Score 41.2; DB 13; Length 588;
Best Local Similarity 56.7%; Pred. No. 2.6;
Matches 76; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 1219 ATTCGGGAACATCTTTTCGCGACCCATCGTAACCGGACGGCGCTCTGCTGTCGAC 1278
|| ||||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 454 ATCTCGGCACATCCATGTCGTGCGCGACGTCGAGCGCTCGCGCGCTGCTCGCGAGC 513

QY 1279 AAATACCGCTGATGAGCAACACACCTGCGTACCACCTGCTGACACGGCTCAGGAC 1338
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 514 GCGACCCGGAGTGGAGCCCGCGCGTGCCTGCGCGCTCATGACGACCGCTACTCC 573

QY 1339 ATCGGTGCAGTCGG 1352
| | | | |
Db 574 ACGTACACCGGCGG 587

RESULT 35
CNS021SP

```

```

LOCUS CNS021SP
DEFINITION Tetraodon nigroviridis genome survey sequence T7 end of clone
226G16 of library G from Tetraodon nigroviridis, genomic survey
sequence.
ACCESSION ALI77298
VERSION 1
KEYWORDS GSS; genome survey sequence.
SOURCE Tetraodon nigroviridis.
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.
REFERENCE 1 (bases 1 to 847)
AUTHORS Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
Saurin,W. and Weissenbach,J.
HUMAN gene number estimate provided by genome wide analysis using
Tetraodon nigroviridis DNA sequence
Unpublished
REFERENCE 2 (bases 1 to 847)
AUTHORS Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,
Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
Weissenbach,J.
TITLE Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 847)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (12-APR-2000)
COMMENT This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetraodon.

FEATURES
source
1..847
Location/Qualifiers
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone="226G16"
/clone_lib="G"
/notes="Genoscope sequence ID : C0AG226BD08LP1-end : T7"

BASE COUNT 199 a 209 c 231 g 203 t 5 others
ORIGIN

Query Match 1.3%; Score 41.2; DB 17; Length 847;
Best Local Similarity 49.1%; Pred. No. 3;
Matches 106; Conservative 1; Mismatches 109; Indels 0; Gaps 0;

QY 2635 ATCGGCTATCTCAAGGCGCTGTCTCTCTAGCGACGTACAAACACAGCATCGCCGACG 2694
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 182 ATCAGCATCGGCATCGGCACCTGCACCTGCAGCAACAGCATCAGCACCAGGACC 241

QY 2695 ACCGGTCGGACGAAACATCGGAAAGGACGCTCAACGGCAGCTGATGCGAGTCGGCGCA 2754
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 242 TGCAGCATTTAGCGCGCATGATGCTAGCATCGTAGGGGGAGGAGGATGGCGCGACAG 301

QY 2755 CTGGGCGGTGTCAACGTTCCGTTTGGCGCAACGGGAGATTGACGGTCGGAAGCGGTCTG 2814
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 302 CTGACCCCGCAATTTTGTGATCTCCAGCAGGCGCATTTATGCMCGTCGACGCGGTGGA 361

QY 2815 CGCTACGACCTGCTCAACAGGATGATTCGCCGAA 2850
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 362 TCGGGGATTTCTGACGAGATGATGATTTCACGCA 397

RESULT 36
LOCUS BG368419
DEFINITION HVSME10018117f Hordeum vulgare 20 DAP spike EST library HVCN0010
(20 DAP) Hordeum vulgare cDNA clone HVSME10018117f, mRNA sequence.
ACCESSION BG368419
VERSION 1
GI:13257520

```



```
QY 2328 TACCGCCGCCATGCGGATATGACGAGACGCCGGCTGAAGACGCGTATCGACGGTTGGA 2387
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 375 GACGACTAGGCAAGAGGCGCCAGTGAGGAGGAAGGAGGCGGTTCAGAGATGTTGA 434
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2388 CCACAACGCTACGGTCTGCGGCTATCGCGGCAACCCCAACAGGACGGTGGAAACGTGGGA 2447
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 435 GGACACCCCTCCCACTCTCGCCCTTCTGCGGACGCGGAGTGGGGGGGTGACATGGGC 494
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2448 ACAGGCGGTGTTGAAGCAAAATGCGCGCAGTACCCCAAC 2489
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 495 CCTAGGCAGACTCAAGCCGACCGAGCAGTGGACTCGAAC 536
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 38
BI790356/c
LOCUS
DEFINITION
  BI790356 631 bp mRNA linear EST 12-MAR-2002
  Id06e08.x1 Melton Normalized Mixed Mouse Pancreas 1 NI-MMS1 Mus
  musculus cDNA clone IMAGE:5662550 3' similar to TR:Q13151 Q13151
  HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN A0 ;, mRNA sequence.
BI790356
BI790356.1 GI:15818081
ACCESSION
VERSION
KEYWORDS
SOURCE
  EST.
  house mouse.
  Mus musculus
REFERENCE
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
  1 (bases 1 to 631)
AUTHORS
  Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
  Lemishka,I., Searce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
  Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
  Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
  ,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T.
  , Jackson,Y. and Bowers,Y.
  Endocrine Pancreas Consortium
  Unpublished (2000)
  Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
  Endocrine Pancreas Consortium
  Harvard University, Howard Hughes Medical Institute
  Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
  MA 02138
  Tel: 617-495-1812
  Fax: 617-495-8557
  Email: dmelton@biohp.harvard.edu
  Library was constructed by Dr. Douglas Melton DNA sequencing by:
  Washington University Genome Sequencing Center For information on
  obtaining a clone please contact: Juliana Brown
  (brownj@fas.harvard.edu)
  MGI:1948876 This sequence now available from the IMAGE consortium,
  for clone orders contact: info@image.llnl.gov
  Seq primer: -40UP from Gibco
  High quality sequence stop: 462.

FEATURES
  source
    1..631
      Location/Qualifiers
        /organism="Mus musculus"
        /strain="ICR"
        /db_xref="taxon:10090"
        /clone="IMAGE:5662550"
        /clone_lib="Melton Normalized Mixed Mouse Pancreas 1
        NI-MMS1"
        /sex="Both for embryonic & newborn, male for adult and
        adult islet"
        /dev_stage="Embryonic day 10.5, E12.5, E16.5, newborn,
        adult, mixed"
        /lab_host="DH10B"
        /note="Vector: pSPORT1; Site_1: Not I; Site_2: Sal I; Five
        libraries representing E10.5/12.5 pancreatic bud, E16.5
        pancreas, newborn pancreas, adult pancreas, and adult
        islets of Langerhans were separately constructed using
        Superscript Plasmid Library kit (Life Technologies). cDNA
        was made by oligo-dT priming and size-selected by column
        fractionation. Libraries were amplified once on solid
        support and plasmid DNA from each library was prepared
```

and mixed in equal amounts. The mixed library DNA was normalized by method #4 from Bonaïdo, Lennon, and Soares 1996 Genome Research 6:791-806; 0.5 microgram single-stranded mixed library plasmid DNA was mixed with 5 micrograms PCR product representing mixed library inserts and hybridized to an EcoT of 6. Single-stranded (unhybridized) plasmids were isolated by hydroxyapatite chromatography and used to make this library."

BASE COUNT 76 a 267 c 185 g 102 t 1 others  
ORIGIN

Query Match 1.3%; Score 40.8; DB 13; Length 631;  
Best Local Similarity 48.3%; Pred. No. 3.5;  
Matches 114; Conservative 0; Mismatches 122; Indels 0; Gaps 0;

```
QY 2968 GCGCTGAACGCGACCTGACGACGCGACTACACGTTACGGCGGCTTACCGGGCG 3027
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 281 GCGCGGAACGACGACGACGAGCGGCTGGCCAGGGCGGCTGGCGGCGCGGCGCTAC 222
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 3028 ACTGCAACCAACCGGACGAGCGGGGACGCAATATGCCGACACCGCCCTGTTGTCGGT 3087
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 221 AACAGCTACGGCGGCTACGGGGGCTACGGGGGCGGTGCGGGCGGCGGCGGCG 162
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 3088 CTGGGCGCGGATGTCGAATTCGGCAACGGCTGGAACGGCTTGGACGCTTACAGCTACGCC 3147
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 161 TCATACGGCGGCGACGACTACGCAACGGCTTCGGCGGCTTCGGCAGCTACAGCCAGCAC 102
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 3148 GGTTCACCAACAGTACGGCAACACGACGGGACGAGTGGCGCTAGGCTACCGGTTCTG 3203
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 101 CAGTCTCTCTACGGGCGGATGAAGAGCGGCGGAGGCGGCGGCGGCGGCGGCGAGCTG 46
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 39  
BM068591  
LOCUS  
DEFINITION  
 BM068591 474 bp mRNA linear EST 13-NOV-2001  
 WHE3459\_D08\_G15ZS Wheat pre-anthesis spike cDNA library Triticum  
 aestivum cDNA clone WHE3459\_D08\_G15, mRNA sequence.  
BM068591  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
 EST.  
 bread wheat.  
 Triticum aestivum

REFERENCE  
 1 (bases 1 to 474)  
 Anderson,O.D., Chao,S., Choi,D.W., Close,T.J., Fenton,R.D., Han  
 ,P.S., Hsia,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J.,  
 Seaton,C.L. and Tong,J.C.  
 The structure and function of the expressed portion of the wheat  
 genomes - Pre-anthesis spike cDNA library

JOURNAL  
 Unpublished (2000)  
COMMENT  
 Contact: Olin Anderson  
 US Department of Agriculture, Agriculture Research Service, Pacific  
 West Area, Western Regional Research Center  
 800 Buchanan Street, Albany, CA 94710, USA  
 Tel: 5105595773  
 Fax: 5105595818  
 Email: oanderson@pw.usda.gov  
 Sequence have been trimmed to remove vector sequence and low  
 quality sequence with phred score less than 20  
 Seq primer: StrataGene SK primer.

FEATURES  
 source  
 /cultivar="Chinese Spring"  
 /db\_xref="taxon:4565"  
 /clone="WHE3459\_D08\_G15"  
 /clone\_lib="Wheat pre-anthesis spike cDNA library"  
 /tissue\_type="Spike before anthesis"  
 /dev\_stage="Adult plant"  
 /lab\_host="E. coli SOLR"  
 /note="Vector: Lambda Uni-ZAP XR, excised phagemid;





GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 27, 2003, 20:44:39 ; Search time 5463 seconds  
(without alignments)  
17068.532 Million cell updates/sec

Title: US-09-830-433A-7  
Perfect score: 3204  
Sequence: 1 atggaacagaccacacatt.....gcgtaggctaccggtctga 3204

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_hgt.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vl.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*
- 29: em\_vl.\*
- 30: em\_hgt\_hum.\*
- 31: em\_hgt\_inv.\*
- 32: em\_hgt\_other.\*
- 33: em\_hgt\_mus.\*
- 34: em\_hgt\_pln.\*
- 35: em\_hgt\_rod.\*
- 36: em\_hgt\_mam.\*
- 37: em\_hgt\_vrt.\*
- 38: em\_sy.\*
- 39: em\_hgt\_hum.\*
- 40: em\_hgt\_mus.\*
- 41: em\_hgt\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	3204	100.0	3204	6	AX024064	Sequence
2	3153	98.4	3204	1	NME311654	AX024064 Sequence
3	2156	67.3	349061	1	NMA311654	AX024064 Sequence
4	834	26.0	4218	6	AX236469	AX236469 Sequence
5	834	26.0	4218	6	AX236469	AX236469 Sequence
6	834	26.0	10591	1	AE002545	AE002545 Neisseria
7	834	26.0	34980	6	AX044034	AX044034 Sequence
8	831	25.9	3939	6	AX236409	AX236409 Sequence
9	831	25.9	3939	6	AX236419	AX236419 Sequence
10	831	25.9	3939	6	AX239717	AX239717 Sequence
11	831	25.9	3939	6	AX239729	AX239729 Sequence
12	831	25.9	4170	6	AX236463	AX236463 Sequence
13	831	25.9	4170	6	AX239749	AX239749 Sequence
14	831	25.9	4179	6	AX236413	AX236413 Sequence
15	831	25.9	4179	6	AX239721	AX239721 Sequence
16	831	25.9	4335	6	AX236457	AX236457 Sequence
17	831	25.9	4335	6	AX239743	AX239743 Sequence
18	831	25.9	4344	6	AX236411	AX236411 Sequence
19	831	25.9	4344	6	AX239719	AX239719 Sequence
20	831	25.9	4425	6	AX236407	AX236407 Sequence
21	831	25.9	4425	6	AX239715	AX239715 Sequence
22	783	24.4	3254	1	NME277537	AX277537 Neisseria
23	321	10.0	600	1	AF169448	AF169448 Neisseria
24	177	5.5	286	1	AF169473	AF169473 Neisseria
25	177	5.5	286	6	A68924	A68924 Sequence 95
26	35	1.1	50	6	AX236815	AX236815 Sequence
27	33	1.0	1434	1	NGOPAK	X52364 N.gonorhoe
28	25	0.8	37497	2	AC017356	AC017356 Drosophila
29	25	0.8	161278	3	AC009749	AC009749 Drosophila
30	25	0.8	327209	3	AE003575	AE003575 Drosophila
31	23	0.7	30	6	AX024142	AX024142 Sequence
32	23	0.7	14043	1	AB014075	AB014075 Clostridi
33	22	0.7	48626	2	AC013119	AC013119 Drosophila
34	22	0.7	85918	2	DMBR33H2	AL121812 Drosophila
35	22	0.7	160986	3	AC108481	AC108481 Drosophila
36	22	0.7	267113	2	AC099761	AC099761 Mus muscu
37	22	0.7	294218	3	AE003430	AE003430 Drosophila
38	21	0.7	34	6	AX024141	AX024141 Sequence
39	21	0.7	348	5	AF281339	AF281339 Oncorhync
40	21	0.7	768	9	HS332365	AF332365 Homo sapi
41	21	0.7	12211	1	AE004895	AE004895 Pseudomon
42	21	0.7	121118	2	AC114017	AC114017 Rattus no
43	21	0.7	124667	2	AC015568	AC015568 Homo sapi
44	20	0.6	735	6	ARI40161	ARI40161 Sequence
45	20	0.6	735	6	ARI99650	ARI99650 Sequence
46	20	0.6	735	6	E16589	E16589 Serratia fi
47	20	0.6	5019	6	AX106914	AX106914 Sequence
48	20	0.6	10034	1	AE002016	AE002016 Deinococc
49	20	0.6	12250	1	AE007837	AE007837 Clostridi
50	20	0.6	20409	2	AC014557	AC014557 Drosophila
51	20	0.6	36328	3	LMFL5213	AL583932 Leishmani
52	20	0.6	89047	6	AX067459	AX067459 Sequence
53	20	0.6	118567	2	AC114134	AC114134 Rattus no
54	20	0.6	127506	8	OSJN00096	AL606647 Oryza sat
55	20	0.6	150962	2	AL627100	AL627100 Homo sapi
56	20	0.6	152763	2	AC119715	AC119715 Rattus no
57	20	0.6	152985	9	AC020581	AC020581 Homo sapi
58	20	0.6	157448	2	AC131139	AC131139 Rattus no
59	20	0.6	159717	2	AC008368	AC008368 Trypanoso
60	20	0.6	159717	2	AC008368	AC008368 Trypanoso
61	20	0.6	160551	9	AC009329	AC009329 Human Chr
62	20	0.6	161252	2	AC127258	AC127258 Mus muscu
63	20	0.6	163115	2	AC101752	AC101752 Mus muscu
64	20	0.6	164942	3	AC011758	AC011758 Drosophila
65	20	0.6	173722	2	AL450993	AL450993 Homo sapi

c 66	20	0.6 175621	2	AC053486	Homo sapi	139	19	0.6 122122	2	AC121621	AC121621 Rattus no
c 67	20	0.6 177182	10	AL669946	Mouse DNA	140	19	0.6 124636	9	AC091690	AC091690 Oryza sat
c 68	20	0.6 180247	2	AC053509	Homo sapi	141	19	0.6 126169	9	AC010291	AC010291 Homo sapi
c 69	20	0.6 185447	9	AL450992	Human DNA	142	19	0.6 127982	9	AC104473	AC104473 Oryza sat
c 70	20	0.6 191053	10	AL672249	Mouse DNA	c 143	19	0.6 132933	9	AL137002	AL137002 Human DNA
71	20	0.6 204702	9	AL356957	Human DNA	144	19	0.6 134764	9	AP003305	AP003305 Homo sapi
72	20	0.6 2122873	3	AC1122873	Mus muscu	c 145	19	0.6 139287	2	AC060794	AC060794 Homo sapi
c 73	20	0.6 216686	3	AE003571	Drosophila	c 146	19	0.6 141073	2	AP005512	AP005512 Oryza sat
c 74	19	0.6 29	6	AX236809	Sequence	147	19	0.6 145456	9	AL513366	AL513366 Human DNA
75	19	0.6 32	32	AX236813	Sequence	c 148	19	0.6 145882	2	AL139131	AL139131 Homo sapi
c 76	19	0.6 126	6	AX017688	Sequence	149	19	0.6 146927	2	AC020144	AC020144 Drosophila
c 77	19	0.6 300	11	GL3553	human STR S	150	19	0.6 150643	2	AC102184	AC102184 Mus muscu
c 78	19	0.6 426	4	AY011990	Sorex ara	151	19	0.6 151024	2	AC016855	AC016855 Homo sapi
c 79	19	0.6 654	9	HSM802875	Sequence	152	19	0.6 151031	9	AC112222	AC112222 Homo sapi
80	19	0.6 807	6	AX119139	Homo sapi	153	19	0.6 151712	9	AC018992	AC018992 Homo sapi
81	19	0.6 1061	8	AF053077	Nicotiana	154	19	0.6 152560	9	AC011339	AC011339 Homo sapi
c 82	19	0.6 1265	9	HMPK027	Human vitam	155	19	0.6 153388	8	AP003289	AP003289 Oryza sat
c 83	19	0.6 1286	6	BD012333	Genes rel	156	19	0.6 153557	2	AC019260	AC019260 Homo sapi
84	19	0.6 1286	3	AK027731	Homo sapi	c 157	19	0.6 155383	2	AP004396	AP004396 Oryza sat
85	19	0.6 1286	23	BD005013	Genes rel	c 158	19	0.6 156551	2	AP003490	AP003490 Oryza sat
86	19	0.6 1296	9	BC012379	Homo sapi	c 159	19	0.6 157722	3	AC092248	AC092248 Drosophila
87	19	0.6 1383	6	AX119157	Sequence	160	19	0.6 158932	2	AC126845	AC126845 Rattus no
88	19	0.6 1451	9	AK000962	Homo sapi	161	19	0.6 159101	2	AC126205	AC126205 Rattus no
c 89	19	0.6 1485	9	HMPK027	Human prote	c 162	19	0.6 159311	2	HS279F22	HS279F22 Homo sapi
c 90	19	0.6 1551	6	AX409708	Sequence	c 163	19	0.6 159734	2	AC118428	AC118428 Rattus no
c 91	19	0.6 1551	9	HMPK027	Human prote	164	19	0.6 162010	2	AP004121	AP004121 Oryza sat
c 92	19	0.6 1596	6	AX119159	Sequence	165	19	0.6 163156	2	AP004085	AP004085 Oryza sat
c 93	19	0.6 1700	10	BC027427	Mus muscu	166	19	0.6 164970	2	AC098136	AC098136 Rattus no
94	19	0.6 1781	6	AX400011	Sequence	167	19	0.6 166512	3	AC099010	AC099010 Drosophila
95	19	0.6 1828	9	AK000382	Homo sapi	168	19	0.6 166727	2	AC130190	AC130190 Pan trogl
c 96	19	0.6 1866	9	BC001246	Homo sapi	c 169	19	0.6 167115	9	AC090093	AC090093 Homo sapi
c 97	19	0.6 2342	9	AB033749	Homo sapi	c 170	19	0.6 168058	3	AC010049	AC010049 Drosophila
c 98	19	0.6 2796	3	AF043944	Mytilus e	171	19	0.6 168267	8	AP003535	AP003535 Oryza sat
99	19	0.6 3050	9	BC024153	Homo sapi	c 172	19	0.6 169414	9	AC068707	AC068707 Homo sapi
100	19	0.6 3127	10	BC004674	Mus muscu	173	19	0.6 169989	9	AC090137	AC090137 Homo sapi
101	19	0.6 3244	9	D50912	Human mRNA	174	19	0.6 170232	9	AL135903	AL135903 Human DNA
102	19	0.6 3301	9	HSM802123	Homo sapi	175	19	0.6 170330	9	AC104430	AC104430 Homo sapi
103	19	0.6 3402	9	AK055116	Homo sapi	176	19	0.6 173184	9	AC112515	AC112515 Homo sapi
104	19	0.6 3418	9	BC004181	Homo sapi	c 177	19	0.6 174345	2	AC025236	AC025236 Homo sapi
105	19	0.6 3423	8	BC008733	Homo sapi	178	19	0.6 174646	2	AC087355	AC087355 Homo sapi
c 106	19	0.6 3889	8	FSY250467	Pinus syl	c 179	19	0.6 174974	3	AC092228	AC092228 Drosophila
107	19	0.6 4195	9	AB058700	Homo sapi	c 180	19	0.6 176123	9	AL390718	AL390718 Human DNA
c 108	19	0.6 7902	2	HUMDNAAPK	L27425 Homo sapien	181	19	0.6 176562	2	AC103686	AC103686 Homo sapi
c 109	19	0.6 8902	2	AC017749	Drosophila	182	19	0.6 178248	9	AC012476	AC012476 Homo sapi
110	19	0.6 10029	1	AE0013261	Methanosa	c 183	19	0.6 180905	2	AC011688	AC011688 Homo sapi
c 111	19	0.6 10670	1	AE009117	Agrobacte	c 184	19	0.6 181004	2	AC116431	AC116431 Papio cyn
c 112	19	0.6 10695	1	AE007265	Sinorhizo	185	19	0.6 181727	2	AC036232	AC036232 Homo sapi
c 113	19	0.6 10871	1	AE004643	Pseudomon	186	19	0.6 183242	9	AC007906	AC007906 Homo sapi
c 114	19	0.6 11516	1	AE008082	Agrobacte	c 187	19	0.6 183970	2	AC017022	AC017022 Homo sapi
c 115	19	0.6 11627	1	AE012368	Xanthomon	c 188	19	0.6 184852	2	AC026032	AC026032 Homo sapi
c 116	19	0.6 16418	9	AF440358	Homo sapi	189	19	0.6 187023	2	AC115422	AC115422 Rattus no
c 117	19	0.6 19482	1	AE007072	Mycobacte	c 190	19	0.6 187023	2	AC115422	AC115422 Rattus no
c 118	19	0.6 30000	6	AX250261	Sequence	191	19	0.6 187533	2	AP004806	AP004806 Oryza sat
c 119	19	0.6 34150	1	MTCY190	270283 Mycobacteri	192	19	0.6 187658	2	AC098041	AC098041 Rattus no
c 120	19	0.6 36348	3	AC024765	Caenorhab	193	19	0.6 189055	2	AC073716	AC073716 Mus muscu
c 121	19	0.6 37918	9	AP005232	Homo sapi	194	19	0.6 191729	2	AL590872	AL590872 Homo sapi
c 122	19	0.6 38080	9	HS429E7	Human DNA	c 195	19	0.6 192001	2	AC019139	AC019139 Homo sapi
123	19	0.6 38141	2	AC018202	Drosophila	c 196	19	0.6 193452	2	AC127575	AC127575 Mus muscu
c 124	19	0.6 40327	9	AP005379	Homo sapi	197	19	0.6 193828	9	AP003025	AP003025 Homo sapi
125	19	0.6 45998	9	HS081031	U81031 Homo sapien	c 198	19	0.6 195640	2	AC094441	AC094441 Rattus no
c 126	19	0.6 62909	6	AX067457	Sequence	c 199	19	0.6 196413	2	AC097269	AC097269 Pan trogl
c 127	19	0.6 69368	2	AC118642	Mus muscu	200	19	0.6 196668	2	AC117584	AC117584 Mus muscu
128	19	0.6 71994	2	AC123649	Mus muscu	201	19	0.6 197063	2	AL807240	AL807240 Mus muscu
c 129	19	0.6 72179	2	AC096066	Rattus no	c 202	19	0.6 200622	9	AC068228	AC068228 Homo sapi
c 130	19	0.6 82849	3	AC004277	Drosophila	203	19	0.6 200799	8	AF137379	AF137379 Nephroset
c 131	19	0.6 86896	1	RCU57682	Rhodobacter	c 204	19	0.6 200799	8	AF137379	AF137379 Nephroset
c 132	19	0.6 89453	8	AC068655	Genomic S	c 205	19	0.6 200990	2	AC099134	AC099134 Rattus no
c 133	19	0.6 93896	8	AB046435	Arabidops	c 206	19	0.6 202457	2	AC121759	AC121759 Homo sapi
c 134	19	0.6 93896	8	F15115	Arabidops	c 207	19	0.6 203732	2	AC129320	AC129320 Mus muscu
c 135	19	0.6 104170	3	AC091510	Leishmani	c 208	19	0.6 204167	2	AC110553	AC110553 Mus muscu
c 136	19	0.6 106763	3	AP002091	Homo sapi	c 209	19	0.6 205050	1	AL646082	AL646082 Ralstonia
c 137	19	0.6 110000	2	LMFICHR36_13	Continuation (14 o	c 210	19	0.6 211544	9	AC025165	AC025165 Homo sapi
c 138	19	0.6 118860	2	AP004084	Oryza sat	c 211	19	0.6 215188	2	AC124196	AC124196 Mus muscu

c 212	19	0.6	216812	10	AL672073	AL672073 Mouse DNA	c 285	18	0.6	1890	8	DCA249962	AJ249962 Daucus ca
c 213	19	0.6	218277	2	AL671741	AL671741 Mus muscu	c 286	18	0.6	1963	9	AF051321	AF051321 Homo sapi
c 214	19	0.6	222845	2	AC094653	AC094653 Rattus no	c 287	18	0.6	2027	9	AX381046	AX381046 Sequence
c 215	19	0.6	227117	2	AC122341	AC122341 Mus muscu	c 288	18	0.6	2080	9	CNS07E50	AL583864 Human chr
c 216	19	0.6	227391	2	AC125160	AC125160 Mus muscu	c 289	18	0.6	2104	8	AB049128	AL583864 Human chr
c 217	19	0.6	237712	9	AC012634	AC012634 Homo sapi	c 290	18	0.6	2160	5	SS17257	Y17257 Scyllorhinu
c 218	19	0.6	237952	2	AL591116	AL591116 Homo sapi	c 291	18	0.6	2218	4	CFA388555	Y17257 Scyllorhinu
c 219	19	0.6	238741	2	AC068894	AC068894 Homo sapi	c 292	18	0.6	2263	9	AK054781	AK054781 Canis fam
c 220	19	0.6	240131	2	AL772361	AL772361 Mus muscu	c 293	18	0.6	2283	5	AF033189	AK054781 Canis fam
c 221	19	0.6	240197	3	AE003643	AE003643 Drosophill	c 294	18	0.6	2363	8	E32353	AF033189 Homo sapi
c 222	19	0.6	242224	2	AC103337	AC103337 Rattus no	c 295	18	0.6	2363	8	E32353	E32353 Gallus ga
c 223	19	0.6	251997	3	AE001218	AE001218 Homo sapi	c 296	18	0.6	2457	3	AF037368	E32353 Plant havin
c 224	19	0.6	269823	2	AE003615	AE003615 Drosophill	c 297	18	0.6	2641	9	AY118678	AF037368 Cucumis m
c 225	19	0.6	270150	9	AE006639	AE006639 Homo sapi	c 298	18	0.6	2659	3	DMBR22	AY118678 Drosophill
c 226	19	0.6	272545	2	AC090533	AC090533 Mus muscu	c 299	18	0.6	2682	9	AK001406	AK091388 Homo sapi
c 227	19	0.6	305502	3	DROSADH03	AE003409 Drosophill	c 300	18	0.6	2837	9	HSCG16P5	X54665 D.melanogas
c 228	19	0.6	305516	3	AE003594	AE003594 Drosophill	c 301	18	0.6	2840	3	DMBRNS23	AK001406 Homo sapi
c 229	19	0.6	309805	2	AC026340	AC026340 Homo sapi	c 302	18	0.6	2867	1	SE2XREC	AK001218 Homo sapi
c 230	19	0.6	338579	1	AP003004	AP003004 Mesorhizo	c 303	18	0.6	2979	9	HS181914	X54664 D.melanogas
c 231	18	0.6	28	6	AX236792	AX236792 Sequence	c 304	18	0.6	3020	10	MUSAPOIVA	M35138 Plasmid pSE
c 232	18	0.6	28	6	AX239766	AX239766 Sequence	c 305	18	0.6	3033	3	AF257595	L81914 Homo sapien
c 233	18	0.6	60	6	AX125920	AX125920 Sequence	c 306	18	0.6	3065	3	AF257606	M13966 Mouse apoli
c 234	18	0.6	60	6	I24287	I24287 Sequence 74	c 307	18	0.6	3068	3	AF257590	AF257595 Drosophill
c 235	18	0.6	111	6	AX079073	AX079073 Sequence	c 308	18	0.6	3068	3	AF257591	AF257606 Drosophill
c 236	18	0.6	150	6	A01406	A01406 Malaria par	c 309	18	0.6	3068	3	AF257598	AF257591 Drosophill
c 237	18	0.6	150	6	A13149	A13149 Nucleotide	c 310	18	0.6	3068	3	AF257599	AF257598 Drosophill
c 238	18	0.6	150	6	A16111	A16111 oligonucleo	c 311	18	0.6	3068	3	AF257600	AF257591 Drosophill
c 239	18	0.6	213	6	AX436315	AX436315 Sequence	c 312	18	0.6	3068	3	AF257601	AF257600 Drosophill
c 240	18	0.6	224	8	AY022882	AY022882 Oryza sat	c 313	18	0.6	3078	3	AF257604	AF257601 Drosophill
c 241	18	0.6	224	8	AY023142	AY023142 Oryza sat	c 314	18	0.6	3078	3	AF257592	AF257604 Drosophill
c 242	18	0.6	231	3	AF272155	AF272155 Procamb	c 315	18	0.6	3078	3	AF257593	AF257592 Drosophill
c 243	18	0.6	328	11	G41605	G41605 Z15424 Zebr	c 316	18	0.6	3078	3	AF257596	AF257593 Drosophill
c 244	18	0.6	558	8	AB036884	AB036884 Arabidops	c 317	18	0.6	3078	3	AF257597	AF257596 Drosophill
c 245	18	0.6	588	6	AX381056	AX381056 Sequence	c 318	18	0.6	3078	3	AF257599	AF257597 Drosophill
c 246	18	0.6	661	6	AX381052	AX381052 Sequence	c 319	18	0.6	3078	3	AF257602	AF257599 Drosophill
c 247	18	0.6	687	9	HS3434262	AJ334262 Homo sapi	c 320	18	0.6	3078	3	AF257603	AF257602 Drosophill
c 248	18	0.6	716	9	HS338190	AJ338190 Homo sapi	c 321	18	0.6	3085	3	AF257607	AF257603 Drosophill
c 249	18	0.6	740	9	HS336538	AJ336538 Homo sapi	c 322	18	0.6	3085	3	AF257608	AF257607 Drosophill
c 250	18	0.6	772	9	HS341989	AJ341989 Homo sapi	c 323	18	0.6	3239	10	BC025581	AF257608 Drosophill
c 251	18	0.6	804	1	AB039533	AB039533 Pseudomon	c 324	18	0.6	3352	8	ATY14404	BC025581 Mus muscu
c 252	18	0.6	837	9	AB049946	AB049946 Homo sapi	c 325	18	0.6	3376	10	BC028975	Y14404 Arabidopsis
c 253	18	0.6	885	6	BC031336	BC031336 Homo sapi	c 326	18	0.6	3476	10	SCYOR378W	BC028975 Mus muscu
c 254	18	0.6	886	9	AF265439	AF265439 Homo sapi	c 327	18	0.6	3647	9	AF151723	Z75286 S.cerevisia
c 255	18	0.6	904	9	AF153607	AF153607 Homo sapi	c 328	18	0.6	3647	9	AF151723	A95922 Sequence 1
c 256	18	0.6	943	9	AF172872	AF172872 Unculture	c 329	18	0.6	3808	3	DMBRQ121	AF151723 Homo sapi
c 257	18	0.6	966	1	AF172915	AF172915 Unculture	c 330	18	0.6	3869	3	DMU51585	X54663 D.melanogas
c 258	18	0.6	968	1	AF172884	AF172884 Unculture	c 331	18	0.6	3916	8	PCH243296	U51585 Drosophila
c 259	18	0.6	972	1	AF172884	AF172884 Unculture	c 332	18	0.6	4003	3	DMERTNT	AJ243296 Penicilli
c 260	18	0.6	1044	3	AX433900	AX433900 Sequence	c 333	18	0.6	4174	3	AY069756	X54666 D.melanogas
c 261	18	0.6	1080	10	AF469004	AF469004 Mus muscu	c 334	18	0.6	4180	6	AX027844	AY069756 Drosophill
c 262	18	0.6	1105	10	AB073819	AB073819 Mus muscu	c 335	18	0.6	4180	6	AX027846	AX027844 Sequence
c 263	18	0.6	1107	4	AF187851	AF187851 Bos tauru	c 336	18	0.6	4180	10	MUSCDB25H	AX027846 Sequence
c 264	18	0.6	1179	1	AB014971	AB014971 Ochrobact	c 337	18	0.6	4338	1	PVU60522	L20899 Mouse cell
c 265	18	0.6	1193	3	AX089425	AX089425 Drosophill	c 338	18	0.6	4479	4	OCFPA1	U60522 Paracoccus
c 266	18	0.6	1281	10	D49684	D49684 Mus muscu	c 339	18	0.6	4509	3	DME271350	X59069 O.cuniculus
c 267	18	0.6	1287	10	MMBP37	X78683 M.musculus	c 340	18	0.6	4509	3	HS803129	AJ271350 Drosophill
c 268	18	0.6	1404	10	AY057054	AY057054 Mus muscu	c 341	18	0.6	4948	9	DM803129	AL81391 Homo sapi
c 269	18	0.6	1530	8	AF009967	AF009967 Candida a	c 342	18	0.6	5166	9	AB040880	AB040880 Homo sapi
c 270	18	0.6	1549	5	AF003532	AF003532 Pleurodel	c 343	18	0.6	5370	1	AF112883	AF112883 Pseudomon
c 271	18	0.6	1557	9	AF069681	AF069681 Homo sapi	c 344	18	0.6	5538	9	AB046773	AB046773 Homo sapi
c 272	18	0.6	1565	4	AF394924	AF394924 Bos tauru	c 345	18	0.6	5710	2	AC013202	AC013202 Drosophill
c 273	18	0.6	1565	4	AF394924	AF394924 Bos tauru	c 346	18	0.6	5788	10	RATMTLL	M74822 Rat MHC cla
c 274	18	0.6	1579	3	AX058394	AX058394 Drosophill	c 347	18	0.6	5865	10	RATMCI	L23127 Rattus norv
c 275	18	0.6	1595	1	MSGDAPDC	M94109 Mycobacteri	c 348	18	0.6	5893	6	AX345761	AX345761 Sequence
c 276	18	0.6	1614	9	AK021626	AK021626 Homo sapi	c 349	18	0.6	5952	6	AX252105	AX252105 Sequence
c 277	18	0.6	1641	9	BC032606	BC032606 Homo sapi	c 350	18	0.6	7167	6	AX345302	AX345302 Sequence
c 278	18	0.6	1645	1	AVU65939	U65939 Azotobacter	c 351	18	0.6	7432	1	PAU81261	U81261 Pseudomonas
c 279	18	0.6	1679	9	AB001835	AB001835 Homo sapi	c 352	18	0.6	7875	6	E43781	E43781 Genes encod
c 280	18	0.6	1689	6	AR129243	AR129243 Sequence	c 353	18	0.6	9299	6	AX346627	AX346627 Sequence
c 281	18	0.6	1691	9	AF055479	AF055479 Homo sapi	c 354	18	0.6	10029	1	AE005292	AE005292 Escherich
c 282	18	0.6	1738	4	BTFT2GEN	X99620 Bos tauru	c 355	18	0.6	10029	1	AE011851	AE011851 Xanthomon
c 283	18	0.6	1807	5	NEWSPA	D63339 Cynops pyr	c 356	18	0.6	10201	1	AE005882	AE005882 Caulobact
c 284	18	0.6	1821	10	RATRMCA	L23128 Rattus norv	c 357	18	0.6	10368	1	AE002071	AE002071 Deinococc

c 358	18	0.6	10487	1	AE009667	AE009667 Brucella	c 431	18	0.6	55739	8	AP000389	AP000389 Arabidops
c 359	18	0.6	10535	1	AE008942	AE008942 Agrobacte	c 432	18	0.6	55892	5	AF013614	AF013614 Fugu rubr
c 360	18	0.6	10575	1	AE009238	AE009238 Agrobacte	c 433	18	0.6	56374	2	AC009855	AC009855 Mus muscu
c 361	18	0.6	10608	1	AE009167	AE009167 Agrobacte	c 434	18	0.6	60009	2	AC023436	AC023436 Homo sapi
c 362	18	0.6	10690	1	AE008133	AE008133 Agrobacte	c 435	18	0.6	60022	2	AC102916	AC102916 Mus muscu
c 363	18	0.6	10723	14	AF204178	AF204178 Dengue vi	c 436	18	0.6	60196	2	AC124052	AC124052 Mus muscu
c 364	18	0.6	10741	1	AE000653	AE000653 Helicobac	c 437	18	0.6	60286	2	AC100131	AC100131 Mus muscu
c 365	18	0.6	10784	1	AE000200	AE000200 Escherich	c 438	18	0.6	61575	2	AC100040	AC100040 Mus muscu
c 366	18	0.6	11067	1	AE011288	AE011288 Chlorobiu	c 439	18	0.6	61575	2	AC100040	AC100040 Mus muscu
c 367	18	0.6	11157	1	AE011938	AE011938 Xanthomon	c 440	18	0.6	62838	2	AC061995	AC061995 Homo sapi
c 368	18	0.6	11176	1	AE007890	AE007890 Agrobacte	c 441	18	0.6	63578	9	AC109356	AC109356 Homo sapi
c 369	18	0.6	11331	1	AE002481	AE002481 Neisseria	c 442	18	0.6	64089	2	AC100165	AC100165 Mus muscu
c 370	18	0.6	11458	1	AE012204	AE012204 Xanthomon	c 443	18	0.6	64295	2	AC100066	AC100066 Mus muscu
c 371	18	0.6	11642	1	AE0005164	AE0005164 Halobacte	c 444	18	0.6	64391	2	AC090856	AC090856 Homo sapi
c 372	18	0.6	11749	1	AE000212	AE000212 Escherich	c 445	18	0.6	64829	2	AC100338	AC100338 Mus muscu
c 373	18	0.6	12250	1	AE012416	AE012416 Xanthomon	c 446	18	0.6	65547	8	AY013245	AY013245 Oryza sat
c 374	18	0.6	12312	1	AE012307	AE012307 Xanthomon	c 447	18	0.6	66281	9	AL138710	AL138710 Human DNA
c 375	18	0.6	12398	1	AE001025	AE001025 Archaeogl	c 448	18	0.6	67555	2	AC100398	AC100398 Mus muscu
c 376	18	0.6	12430	1	AE000370	AE000370 Escherich	c 449	18	0.6	67859	2	AC101982	AC101982 Mus muscu
c 377	18	0.6	12438	1	AE000404	AE000404 Xylella f	c 450	18	0.6	68904	2	AC027370	AC027370 Homo sapi
c 378	18	0.6	12986	6	AX251307	AX251307 Sequence	c 451	18	0.6	71821	2	AC091584	AC091584 Homo sapi
c 379	18	0.6	13086	1	SLG2	AL136500 Streptomy	c 452	18	0.6	72924	2	AC025930	AC025930 Homo sapi
c 380	18	0.6	13223	1	AE004885	AE004885 Pseudomon	c 453	18	0.6	76179	2	AC012976	AC012976 Drosophil
c 381	18	0.6	13949	1	AE005798	AE005798 Caulobact	c 454	18	0.6	76529	8	H0423H10	AL442112 Oryza sat
c 382	18	0.6	13973	8	ATY14403	Y14403 Arabidops	c 455	18	0.6	76606	9	AL162588	AL162588 Human DNA
c 383	18	0.6	14031	1	AE008376	AE008376 Agrobacte	c 456	18	0.6	78276	3	AC024761	AC024761 Caenorhab
c 384	18	0.6	14106	1	AE007007	AE007007 Mycobacte	c 457	18	0.6	80308	2	AC105512	AC105512 Rattus no
c 385	18	0.6	14212	2	OSJN00250	AL731607 Oryza sat	c 458	18	0.6	80622	2	AC109072	AC109072 Rattus no
c 386	18	0.6	14433	1	U67533	U67533 Methanococc	c 459	18	0.6	82912	3	AC004574	AC004574 Drosophil
c 387	18	0.6	14542	6	AX251844	AX251844 Sequence	c 460	18	0.6	83013	2	AC111418	AC111418 Rattus no
c 388	18	0.6	14542	6	AX344230	AX344230 Sequence	c 461	18	0.6	84243	2	AC122115	AC122115 Homo sapi
c 389	18	0.6	14542	6	AX348623	AX348623 Sequence	c 462	18	0.6	84514	10	AL627264	AL627264 Mouse DNA
c 390	18	0.6	14789	1	AE004916	AE004916 Pseudomon	c 463	18	0.6	84702	8	AB018114	AB018114 Arabidops
c 391	18	0.6	16255	1	D90747	D90747 Escherichia	c 464	18	0.6	85995	2	AC027404	AC027404 Homo sapi
c 392	18	0.6	16259	1	R065510	U65510 Rhodospiril	c 465	18	0.6	86530	2	AC111753	AC111753 Rattus no
c 393	18	0.6	16515	1	AB073929	AB073929 Sporosarc	c 466	18	0.6	88318	8	AF002842	AF002842 Oryza sat
c 394	18	0.6	16963	8	SPAC16A10	Z97185 S. pombe chr	c 467	18	0.6	89337	2	AC097874	AC097874 Rattus no
c 395	18	0.6	17212	1	D90746	D90746 Escherichia	c 468	18	0.6	89818	9	AC002126	AC002126 Homo sapi
c 396	18	0.6	17893	6	AX346267	AX346267 Sequence	c 469	18	0.6	91040	2	AC017292	AC017292 Drosophil
c 397	18	0.6	18882	1	D90735	D90735 Escherichia	c 470	18	0.6	91573	9	AC012372	AC012372 Homo sapi
c 398	18	0.6	19001	3	CEM01A8	Z27081 Caenorhabdi	c 471	18	0.6	93234	8	AC009465	AC009465 Arabidops
c 399	18	0.6	19478	3	CEM05E8	Z73975 Caenorhabdi	c 472	18	0.6	93398	2	AF001863	AF001863 Homo sapi
c 400	18	0.6	19650	1	D90736	D90736 Escherichia	c 473	18	0.6	95009	8	ATF25G13	AL079349 Arabidops
c 401	18	0.6	22353	2	AC017269	AC017269 Drosophil	c 474	18	0.6	95129	9	AC003099	AC003099 Homo sapi
c 402	18	0.6	23821	6	E43780	E43780 Genes encod	c 475	18	0.6	95597	2	AC014022	AC014022 Drosophil
c 403	18	0.6	25998	9	AF172080	AF172080 Homo sapi	c 476	18	0.6	95746	2	AC020248	AC020248 Drosophil
c 404	18	0.6	26238	3	U55366	U55366 Caenorhabdi	c 477	18	0.6	96006	2	AL139224	AL139224 Homo sapi
c 405	18	0.6	29227	3	AF000192	AF000192 Caenorhabdi	c 478	18	0.6	96006	9	HSJ590P13	AL109743 Human DNA
c 406	18	0.6	29314	9	AC002455	AC002455 Human cos	c 479	18	0.6	96485	2	AC130635	AC130635 Rattus no
c 407	18	0.6	29589	8	SPAC22F8	AL109831 S. pombe c	c 480	18	0.6	96705	2	AC098198	AC098198 Rattus no
c 408	18	0.6	35516	1	MTCY373	Z73419 Mycobacteri	c 481	18	0.6	96745	2	AC074219	AC074219 Mus muscu
c 409	18	0.6	35594	2	AL589656	AL589656 Human DNA	c 482	18	0.6	97033	2	AC015146	AC015146 Drosophil
c 410	18	0.6	36236	9	AC1589656	AL589656 Human DNA	c 483	18	0.6	97943	9	HSAC002467	AC002467 Human BAC
c 411	18	0.6	36985	1	MSGB1529CS	L78824 Mycobacteri	c 484	18	0.6	98062	2	RMS16010	AL603719 Rattus no
c 412	18	0.6	37578	3	LMFL7171	AL133435 Leishmani	c 485	18	0.6	99504	2	AC098401	AC098401 Rattus no
c 413	18	0.6	38525	3	AF003145	AF003145 Caenorhab	c 486	18	0.6	99802	9	AL158169	AL158169 Human DNA
c 414	18	0.6	38551	5	TRU414048	AJ414048 Takifugu	c 487	18	0.6	99940	9	AC079246	AC079246 Homo sapi
c 415	18	0.6	39582	3	DMC17A9	AL009146 Drosophil	c 488	18	0.6	100786	2	AC013972	AC013972 Drosophil
c 416	18	0.6	39708	3	CEP35G12	Z46242 Caenorhabdi	c 489	18	0.6	102378	2	AP003736	AP003736 Oryza sat
c 417	18	0.6	41349	3	AC0018276	AC0018276 Drosophil	c 490	18	0.6	104577	5	AL713869	AL713869 Zebrafish
c 418	18	0.6	42079	3	AC005941	AC005941 Leishmani	c 491	18	0.6	104726	9	AC008087	AC008087 Homo sapi
c 419	18	0.6	45125	3	AF068714	AF068714 Caenorhab	c 492	18	0.6	105362	9	AC012370	AC012370 Homo sapi
c 420	18	0.6	45597	9	AF000533	AP000533 Homo sapi	c 493	18	0.6	107037	9	HSJ678E16	AL122004 Human DNA
c 421	18	0.6	46533	2	AC122120	AC122120 Mus muscu	c 494	18	0.6	108322	2	AC099662	AC099662 Rattus no
c 422	18	0.6	48622	8	AB008266	AB008266 Arabidops	c 495	18	0.6	108388	9	AP001821	AP001821 Homo sapi
c 423	18	0.6	49600	8	AP004342	AP004342 Oryza sat	c 496	18	0.6	108964	8	AC073391	AC073391 Oryza sat
c 424	18	0.6	50960	2	AC017313	AC017313 Drosophil	c 497	18	0.6	110000	2	LMFLCHR32_19	Continuation (20 o
c 425	18	0.6	51111	2	AC100579	AC100579 Mus muscu	c 498	18	0.6	110000	2	LMFLCHR32_20	Continuation (21 o
c 426	18	0.6	51661	8	AP003864	AP003864 Oryza sat	c 499	18	0.6	110000	2	TBFLCHR34_14	Continuation (15 o
c 427	18	0.6	52586	9	AF258547	AF258547 Homo sapi	c 500	18	0.6	110000	2	TBFLCHR34_14	Continuation (3 of
c 428	18	0.6	53931	2	AC100941	AC100941 Mus muscu	c 501	18	0.6	110000	3	AC125735_0	AC125735 Leishmani
c 429	18	0.6	54299	2	AC124645	AC124645 Mus muscu	c 502	18	0.6	110000	3	AC125735_1	Continuation (2 of
c 430	18	0.6	55175	1	ECU28375	U28375 Escherichia	c 503	18	0.6	110338	3	AC006066	AC006066 Drosophil

c 504	18	0.6 110611	8	AC022520	Arabidops	AC022520	577	18	0.6 143969	8	AP002839	Oryza sat
c 505	18	0.6 110721	2	AC098764	Rattus no	AC098764	578	18	0.6 144191	2	AC079874	Oryza sat
c 506	18	0.6 111172	2	AP005293	Oryza sat	AP005293	c 579	18	0.6 144321	8	AC130603	Oryza sat
c 507	18	0.6 111469	9	AC022470	Homo sapi	AC022470	580	18	0.6 144538	9	CNS07603	Human chr
c 508	18	0.6 111634	9	AC025275	Homo sapi	AC025275	581	18	0.6 144784	9	AC009224	Homo sapi
c 509	18	0.6 111837	9	AC069383	Homo sapi	AC069383	582	18	0.6 144983	2	AC105284	Homo sapi
c 510	18	0.6 113031	9	AL445672	Human DNA	AL445672	c 583	18	0.6 145242	8	AC013457	Homo sapi
c 511	18	0.6 113639	9	HSJ744A17	Human DNA	AL096862	c 584	18	0.6 145299	10	AL603826	Mouse DNA
c 512	18	0.6 113646	2	AC106265	Rattus no	AC106265	585	18	0.6 148370	10	AC091514	Rattus no
c 513	18	0.6 113988	9	AL133546	Human DNA	AL133546	586	18	0.6 148375	2	AC118789	Rattus no
c 514	18	0.6 114070	2	AP005243	Oryza sat	AP005243	c 587	18	0.6 148381	8	AP003253	Oryza sat
c 515	18	0.6 114584	2	AF252827	Homo sapi	AF252827	588	18	0.6 148507	2	AC011263	Homo sapi
c 516	18	0.6 116004	9	AF258545	Homo sapi	AF258545	589	18	0.6 148704	8	AC124213	Genomic s
c 517	18	0.6 116801	2	AL390054	Homo sapi	AL390054	c 590	18	0.6 148750	9	AL596225	Human DNA
c 518	18	0.6 117716	2	AC094474	Rattus no	AC094474	c 591	18	0.6 149347	8	AC073392	Oryza sat
c 519	18	0.6 117779	10	AL732293	Mouse DNA	AL732293	c 592	18	0.6 149945	2	AC124146	Rattus no
c 520	18	0.6 119061	2	AL669909	Mus muscu	AL669909	593	18	0.6 150113	9	AC015988	Homo sapi
c 521	18	0.6 119066	2	AC026377	Mus muscu	AC026377	594	18	0.6 150381	8	AP003794	Oryza sat
c 522	18	0.6 119243	2	AC112576	Rattus no	AC112576	c 595	18	0.6 150690	9	AC090402	Homo sapi
c 523	18	0.6 119295	2	AC127427	Magnaport	AC127427	596	18	0.6 151243	2	AC094894	Homo sapi
c 524	18	0.6 119797	2	AC013236	Drosophil	AC013236	597	18	0.6 151551	2	AC021291	Rattus no
c 525	18	0.6 120030	9	AC002075	Homo sapi	AC002075	c 598	18	0.6 151847	2	AC022792	Homo sapi
c 526	18	0.6 120159	2	AC120509	Oryza sat	AC120509	599	18	0.6 152668	2	AC087623	Homo sapi
c 527	18	0.6 120641	2	AC094693	Rattus no	AC094693	c 600	18	0.6 153297	2	AC099458	Rattus no
c 528	18	0.6 121019	2	AC102708	Mus muscu	AC102708	601	18	0.6 153448	2	AC099458	Rattus no
c 529	18	0.6 122648	8	AC026025	Homo sapi	AC026025	c 602	18	0.6 153566	8	AP003287	Oryza sat
c 530	18	0.6 123427	8	AC091732	Oryza sat	AC091732	603	18	0.6 153789	2	AP004812	Oryza sat
c 531	18	0.6 124093	2	AP005575	Oryza sat	AP005575	c 604	18	0.6 153954	9	AL354862	Human DNA
c 532	18	0.6 124989	2	CNS08087	Genomic s	AC008007	c 605	18	0.6 154242	2	AC073373	Mus muscu
c 533	18	0.6 126155	2	AC125631	Rattus no	AC125631	606	18	0.6 154613	2	AC128712	Homo sapi
c 534	18	0.6 126175	2	AC127717	Rattus no	AC127717	c 607	18	0.6 154641	10	AC121586	Mus muscu
c 535	18	0.6 126315	2	AP004116	Oryza sat	AP004116	c 608	18	0.6 154832	2	AC119142	Rattus no
c 536	18	0.6 126659	8	OSJN00015	Oryza sat	AL606455	c 609	18	0.6 155060	2	AP005591	Oryza sat
c 537	18	0.6 127204	2	AC122114	Atelerix	AC122114	610	18	0.6 155152	2	AP004858	Oryza sat
c 538	18	0.6 127423	2	AC117018	Rattus no	AC117018	c 611	18	0.6 155276	2	AC018410	Homo sapi
c 539	18	0.6 127432	2	AC129056	Rattus no	AC129056	612	18	0.6 156127	2	AP005156	Oryza sat
c 540	18	0.6 128210	2	AP005643	Oryza sat	AP005643	613	18	0.6 156143	2	AC128712	Homo sapi
c 541	18	0.6 128342	9	AL627313	Human DNA	AL627313	c 614	18	0.6 156440	2	AC111368	Rattus no
c 542	18	0.6 129104	2	CNS08092	Genomic s	AC008007	c 615	18	0.6 156472	2	AL732597	Mus muscu
c 543	18	0.6 130235	8	AC008007	Genomic s	AC008007	c 616	18	0.6 156763	9	AC005901	Homo sapi
c 544	18	0.6 130266	9	AL596220	Human DNA	AL596220	c 617	18	0.6 157021	2	AC129066	Bidelphis
c 545	18	0.6 130941	2	AC105770	Oryza sat	AC105770	618	18	0.6 157451	2	AC120275	Rattus no
c 546	18	0.6 132104	2	AC073672	Mus muscu	AC073672	c 619	18	0.6 157926	2	AC096047	Rattus no
c 547	18	0.6 132384	9	AC116347	Homo sapi	AC116347	c 620	18	0.6 158105	2	AC127408	Rattus no
c 548	18	0.6 132755	9	AC008670	Homo sapi	AC008670	c 621	18	0.6 158633	2	AC015472	Homo sapi
c 549	18	0.6 132927	8	AC079890	Oryza sat	AC079890	c 622	18	0.6 158641	2	AC102786	Mus muscu
c 550	18	0.6 133028	8	AP003447	Oryza sat	AP003447	c 623	18	0.6 159191	2	AP005099	Oryza sat
c 551	18	0.6 133461	2	AC025968	Homo sapi	AC025968	c 624	18	0.6 159538	2	AC129676	Homo sapi
c 552	18	0.6 134617	2	AC115001	Mus muscu	AC115001	c 625	18	0.6 159611	9	AC026475	Homo sapi
c 553	18	0.6 134725	2	AC114389	Rattus no	AC114389	c 626	18	0.6 159700	2	AC010822	Homo sapi
c 554	18	0.6 135357	2	AP004769	Oryza sat	AP004769	c 627	18	0.6 159712	9	AP004219	Homo sapi
c 555	18	0.6 135743	2	AC123302	Rattus no	AC123302	c 628	18	0.6 159784	2	AC061996	Homo sapi
c 556	18	0.6 136356	9	AC040166	Homo sapi	AC040166	c 629	18	0.6 159942	9	AC025018	Homo sapi
c 557	18	0.6 136906	2	RN86120	Rattus no	AC040166	630	18	0.6 160026	2	AC021437	Homo sapi
c 558	18	0.6 137071	2	AC119540	Rattus no	AL603805	631	18	0.6 160030	2	CNS07YPB	Oryza sat
c 559	18	0.6 138633	9	AP003078	Homo sapi	AC119540	c 632	18	0.6 160531	3	AC008320	Drosophil
c 560	18	0.6 138893	2	AC107593	Rattus no	AP003078	c 633	18	0.6 160573	9	AC119676	Homo sapi
c 561	18	0.6 138904	2	AP003928	Oryza sat	AC107593	c 634	18	0.6 160635	2	AC129997	Rattus no
c 562	18	0.6 138939	9	AC119675	Homo sapi	AC119675	c 635	18	0.6 160796	9	AL499604	Human DNA
c 563	18	0.6 139005	2	AF286112	Tetraodon	AF286112	636	18	0.6 160831	2	AC025015	Homo sapi
c 564	18	0.6 139468	8	AC090441	Oryza sat	AC090441	c 637	18	0.6 161313	2	AC107846	Mus muscu
c 565	18	0.6 139581	2	AP005055	Oryza sat	AP005055	c 638	18	0.6 161428	8	AC126925	Canis fam
c 566	18	0.6 139814	2	AP003964	Rattus no	AP003964	c 639	18	0.6 161873	8	AC090056	Oryza sat
c 567	18	0.6 140191	2	AC109434	Rattus no	AC109434	c 640	18	0.6 161976	2	AC012415	Homo sapi
c 568	18	0.6 140474	2	AC116947	Oryza sat	AC116947	c 641	18	0.6 162060	2	AL358874	Homo sapi
c 569	18	0.6 140583	2	AP005413	Rattus no	AP005413	c 642	18	0.6 162560	2	AL358874	Homo sapi
c 570	18	0.6 140714	2	RN374E16	Rattus no	AL603726	c 643	18	0.6 162845	2	AC101829	Mus muscu
c 571	18	0.6 140774	2	AC108904	Felis cat	AC108904	c 644	18	0.6 163158	2	AC111278	Rattus no
c 572	18	0.6 141386	2	AC119454	Rattus no	AC119454	c 645	18	0.6 164018	9	AL354868	Human DNA
c 573	18	0.6 141777	9	AC0011738	Homo sapi	AC0011738	c 646	18	0.6 164810	2	AC115069	Mus muscu
c 574	18	0.6 142164	2	AC115193	Rattus no	AC115193	c 647	18	0.6 164911	2	AC119581	Rattus no
c 575	18	0.6 142907	2	AC091330	Trypanoso	AC091330	c 648	18	0.6 165341	2	AC092037	Homo sapi
c 576	18	0.6 143252	2	CNS080CB5	Oryza sat	AL845434	649	18	0.6 165459	2	AC121653	Rattus no

c 650	18	0.6 165901	9	AP003474	AP003474 Homo sapi	723	18	0.6 177320	9	AC104441	AC104441 Homo sapi
c 651	18	0.6 166013	2	AC016328	AC016328 Homo sapi	c 724	18	0.6 177688	9	CNS01DT5	AL132640 Human chr
c 652	18	0.6 166133	2	AC016907	AC016907 Rattus no	c 725	18	0.6 177782	9	AC117532	AC117532 Homo sapi
c 653	18	0.6 166352	3	AC010069	AC010069 Drosophill	c 726	18	0.6 177865	9	AC098483	AC098483 Homo sapi
c 654	18	0.6 166543	3	AC005813	AC005813 Drosophill	c 727	18	0.6 178199	2	AC018772	AC018772 Homo sapi
c 655	18	0.6 166569	3	AC0106496	AC0106496 Rattus no	c 728	18	0.6 178714	2	AC017008	AC017008 Homo sapi
c 656	18	0.6 166592	2	AC126909	AC126909 Rattus no	c 729	18	0.6 178959	2	AP004809	AP004809 Oryza sat
c 657	18	0.6 166899	8	AP003350	AP003350 Oryza sat	c 730	18	0.6 179343	2	AC121904	AC121904 Mus muscu
c 658	18	0.6 167102	9	AC084017	AC084017 Homo sapi	c 731	18	0.6 179683	2	AC101721	AC101721 Mus muscu
c 659	18	0.6 167289	2	AC093187	AC093187 Papio cyn	c 732	18	0.6 179683	2	AC101721	AC101721 Mus muscu
c 660	18	0.6 167617	2	AC091403	AC091403 Sus scrof	c 733	18	0.6 179796	2	AC098948	AC098948 Rattus no
c 661	18	0.6 167630	2	AC105064	AC105064 Mus muscu	c 734	18	0.6 179950	2	AC111288	AC111288 Rattus no
c 662	18	0.6 167694	9	AC113137	AC113137 Homo sapi	c 735	18	0.6 180009	2	AC023259	AC023259 Homo sapi
c 663	18	0.6 167888	2	AC129805	AC129805 Homo sapi	c 736	18	0.6 180073	2	AC068669	AC068669 Homo sapi
c 664	18	0.6 167914	2	AC019268	AC019268 Homo sapi	c 737	18	0.6 180773	2	AC092969	AC092969 Homo sapi
c 665	18	0.6 168187	2	AC094137	AC094137 Rattus no	c 738	18	0.6 180890	2	AC018402	AC018402 Homo sapi
c 666	18	0.6 168505	2	AC107142	AC107142 Rattus no	c 739	18	0.6 180915	2	AC119042	AC119042 Homo sapi
c 667	18	0.6 168562	2	AC107012	AC107012 Rattus no	c 740	18	0.6 181009	3	AC007575	AC007575 Drosophill
c 668	18	0.6 169249	2	AC022815	AC022815 Homo sapi	c 741	18	0.6 181486	2	AC116503	AC116503 Mus muscu
c 669	18	0.6 170025	2	OSJN00031	AL606590 Oryza sat	c 742	18	0.6 181607	9	AC021581	AC021581 Homo sapi
c 670	18	0.6 170082	2	AC129774	AC129774 Mus muscu	c 743	18	0.6 181780	2	AC116814	AC116814 Mus muscu
c 671	18	0.6 170341	3	AC013423	AC013423 Drosophill	c 744	18	0.6 181782	2	AC117172	AC117172 Rattus no
c 672	18	0.6 170369	2	AF214633	AF214633 Homo sapi	c 745	18	0.6 181927	3	AC010573	AC010573 Drosophill
c 673	18	0.6 170523	2	AC095326	AC095326 Homo sapi	c 746	18	0.6 182322	9	AC092379	AC092379 Homo sapi
c 674	18	0.6 170931	2	AC096400	AC096400 Rattus no	c 747	18	0.6 182339	2	AC092761	AC092761 Papio cyn
c 675	18	0.6 170959	2	AC022197	AC022197 Homo sapi	c 748	18	0.6 182528	9	AC020743	AC020743 Homo sapi
c 676	18	0.6 171076	2	AC025560	AC025560 Homo sapi	c 749	18	0.6 182540	2	AC016916	AC016916 Homo sapi
c 677	18	0.6 171114	9	AC092757	AC092757 Homo sapi	c 750	18	0.6 182622	2	AC117173	AC117173 Rattus no
c 678	18	0.6 171168	2	AP005385	AP005385 Oryza sat	c 751	18	0.6 183043	2	AC106096	AC106096 Rattus no
c 679	18	0.6 171179	2	AL162722	AL162722 Homo sapi	c 752	18	0.6 183067	9	AC040914	AC040914 Homo sapi
c 680	18	0.6 171427	9	AC021590	AC021590 Homo sapi	c 753	18	0.6 183193	2	AL844554	AL844554 Mus muscu
c 681	18	0.6 171710	2	AC068049	AC068049 Homo sapi	c 754	18	0.6 183249	9	AC004828	AC004828 Homo sapi
c 682	18	0.6 172024	10	AL606904	AL606904 Mouse DNA	c 755	18	0.6 183338	2	AC069514	AC069514 Homo sapi
c 683	18	0.6 172078	2	AC019359	AC019359 Homo sapi	c 756	18	0.6 183370	3	AC008213	AC008213 Drosophill
c 684	18	0.6 172154	2	AC018396	AC018396 Homo sapi	c 757	18	0.6 183719	2	AC068842	AC068842 Homo sapi
c 685	18	0.6 172277	2	AC099233	AC099233 Rattus no	c 758	18	0.6 184172	2	AC130589	AC130589 Rattus no
c 686	18	0.6 172279	2	AC107529	AC107529 Rattus no	c 759	18	0.6 184267	2	AC078796	AC078796 Homo sapi
c 687	18	0.6 172349	9	AC019185	AC019185 Homo sapi	c 760	18	0.6 184315	2	AP005621	AP005621 Oryza sat
c 688	18	0.6 172367	2	AC118989	AC118989 Canis fam	c 761	18	0.6 184509	2	AC113237	AC113237 Canis fam
c 689	18	0.6 172512	9	AC016903	AC016903 Homo sapi	c 762	18	0.6 184582	9	AC010207	AC010207 Homo sapi
c 690	18	0.6 173645	2	AC123526	AC123526 Oryza sat	c 763	18	0.6 184888	2	AP004795	AP004795 Oryza sat
c 691	18	0.6 173691	2	AC087462	AC087462 Homo sapi	c 764	18	0.6 185096	9	AL355140	AL355140 Human DNA
c 692	18	0.6 173769	9	AC113933	AC113933 Homo sapi	c 765	18	0.6 185165	9	AC008732	AC008732 Homo sapi
c 693	18	0.6 173797	3	AC105898	AC105898 Drosophill	c 766	18	0.6 185311	9	AL355586	AL355586 Human DNA
c 694	18	0.6 173932	2	AC102837	AC102837 Mus muscu	c 767	18	0.6 185563	2	AC112284	AC112284 Rattus no
c 695	18	0.6 174247	2	AC095090	AC095090 Rattus no	c 768	18	0.6 185624	9	AC108162	AC108162 Homo sapi
c 696	18	0.6 174380	2	AP003720	AP003720 Homo sapi	c 769	18	0.6 185937	9	AC107896	AC107896 Homo sapi
c 697	18	0.6 174419	2	AC119060	AC119060 Bos tauru	c 770	18	0.6 185948	2	AC008901	AC008901 Rattus no
c 698	18	0.6 174480	2	AC026069	AC026069 Homo sapi	c 771	18	0.6 185952	9	AC005906	AC005906 Homo sapi
c 699	18	0.6 174602	2	AC106552	AC106552 Rattus no	c 772	18	0.6 186359	2	AC111041	AC111041 Mus muscu
c 700	18	0.6 174616	3	AC010016	AC010016 Drosophill	c 773	18	0.6 186416	2	AC011226	AC011226 Homo sapi
c 701	18	0.6 174793	9	CNS05PE5	AL358334 Human chr	c 774	18	0.6 186714	3	AC023688	AC023688 Drosophill
c 702	18	0.6 174877	2	AP004818	AP004818 Oryza sat	c 775	18	0.6 186973	2	AC124932	AC124932 Rattus no
c 703	18	0.6 174919	2	AC120741	AC120741 Rattus no	c 776	18	0.6 187193	9	AC007222	AC007222 Homo sapi
c 704	18	0.6 175115	3	AC105899	AC105899 Drosophill	c 777	18	0.6 187600	2	AC092515	AC092515 Papio cyn
c 705	18	0.6 175269	2	AC120762	AC120762 Rattus no	c 778	18	0.6 187734	9	AC025918	AC025918 Homo sapi
c 706	18	0.6 175430	2	AC082646	AC082646 Homo sapi	c 779	18	0.6 187908	9	HS495010	AL031121 Human DNA
c 707	18	0.6 175456	2	AC118909	AC118909 Rattus no	c 780	18	0.6 188072	2	AC009040	AC009040 Homo sapi
c 708	18	0.6 175911	2	AC133355	AC133355 Rattus no	c 781	18	0.6 188494	2	AC126154	AC126154 Rattus no
c 709	18	0.6 175928	2	AC129750	AC129750 Rattus no	c 782	18	0.6 188509	2	AC115933	AC115933 Mus muscu
c 710	18	0.6 175947	8	AP003435	AP003435 Oryza sat	c 783	18	0.6 188637	2	AL845306	AL845306 Danio rer
c 711	18	0.6 176010	2	AC069125	AC069125 Homo sapi	c 784	18	0.6 188716	2	AC015886	AC015886 Mus muscu
c 712	18	0.6 176095	2	AC019346	AC019346 Homo sapi	c 785	18	0.6 188740	2	AC024045	AC024045 Homo sapi
c 713	18	0.6 176194	9	AL691486	AL691486 Human DNA	c 786	18	0.6 189008	2	AP001491	AP001491 Homo sapi
c 714	18	0.6 176580	2	AC102248	AC102248 Mus muscu	c 787	18	0.6 189179	2	AC090265	AC090265 Homo sapi
c 715	18	0.6 176580	2	AP003634	AP003634 Oryza sat	c 788	18	0.6 189466	9	AC018887	AC018887 Homo sapi
c 716	18	0.6 176642	10	AC121987	AC121987 Mus muscu	c 789	18	0.6 189722	9	AC067793	AC067793 Homo sapi
c 717	18	0.6 176915	2	AL844176	AL844176 Mus muscu	c 790	18	0.6 189771	2	AC112921	AC112921 Homo sapi
c 718	18	0.6 176928	9	AC084239	AC084239 Homo sapi	c 791	18	0.6 190017	9	AC090589	AC090589 Homo sapi
c 719	18	0.6 176971	2	AC027455	AC027455 Homo sapi	c 792	18	0.6 190091	2	AC123953	AC123953 Mus muscu
c 720	18	0.6 176995	9	AC022735	AC022735 Homo sapi	c 793	18	0.6 191028	8	PFU38804	U38804 Porphyra pu
c 721	18	0.6 177018	2	AC106357	AC106357 Rattus no	c 794	18	0.6 191119	9	AC009957	AC009957 Homo sapi
c 722	18	0.6 177096	3	AC107326	AC107326 Drosophill	c 795	18	0.6 191556	2	AC026117	AC026117 Homo sapi

796	18	0.6 191625	2	AC118038	AC118038 Mus muscu	869	18	0.6 220633	9	HUU91321	U91321 Human Chrom
c 797	18	0.6 191957	9	AC015845	AC015845 Homo sapi	c 870	18	0.6 221883	2	AC104752	AC104752 Mus muscu
798	18	0.6 192321	9	AC119324	AC119324 Rattus no	871	18	0.6 221912	10	AL672003	AL672003 Mouse DNA
c 799	18	0.6 192387	9	CNS01D00	AL133444 Human chr	c 872	18	0.6 222618	2	AC103335	AC103335 Rattus no
c 800	18	0.6 192966	9	AL353813	AL353813 Human DNA	c 873	18	0.6 223280	2	AC130456	AC130456 Homo sapi
801	18	0.6 192968	9	AC099655	AC099655 Rattus no	c 874	18	0.6 223326	2	AC127367	AC127367 Mus muscu
c 802	18	0.6 193046	2	AC026034	AC026034 Homo sapi	875	18	0.6 223898	2	AC124604	AC124604 Mus muscu
c 803	18	0.6 193260	2	AC025598	AC025598 Homo sapi	876	18	0.6 224068	9	CNS01D0UJ	AL133246 BAC sequ
c 804	18	0.6 194329	2	AC116404	AC116404 Mus muscu	c 877	18	0.6 224757	2	AC122913	AC122913 Mus muscu
805	18	0.6 194646	2	AC090129	AC090129 Homo sapi	c 878	18	0.6 224799	2	AC122472	AC122472 Mus muscu
c 806	18	0.6 194790	3	AC009840	AC009840 Drosophill	879	18	0.6 225083	10	AL596382	AL596382 Mouse DNA
807	18	0.6 194905	2	AL603764	AL603764 Homo sapi	880	18	0.6 225136	2	AC122866	AC122866 Mus muscu
808	18	0.6 195270	9	AC096534	AC096534 Homo sapi	881	18	0.6 227538	10	AC002397	AC002397 Mouse chr
809	18	0.6 195296	2	AC026606	AC026606 Homo sapi	c 882	18	0.6 228263	9	AC079930	AC079930 Homo sapi
810	18	0.6 195348	2	AC127081	AC127081 Rattus no	883	18	0.6 228516	9	AC002523	AC002523 Homo sapi
c 811	18	0.6 196259	2	AL135915	AL135915 Homo sapi	884	18	0.6 230458	2	AL731828	AL731828 Mus muscu
c 812	18	0.6 196544	9	AC015974	AC015974 Homo sapi	885	18	0.6 232904	2	AL732346	AL732346 Mus muscu
813	18	0.6 196726	2	AC126723	AC126723 Rattus no	886	18	0.6 232990	2	AL732546	AL732546 Mus muscu
c 814	18	0.6 197078	2	AC117382	AC117382 Homo sapi	887	18	0.6 234189	2	AL125040	AL125040 Mus muscu
c 815	18	0.6 197127	2	AC098253	AC098253 Rattus no	888	18	0.6 238016	2	AL773538	AL773538 Mus muscu
c 816	18	0.6 197437	2	AL772303	AL772303 Mus muscu	c 889	18	0.6 240871	2	AC122568	AC122568 Rattus no
817	18	0.6 197597	3	AC011253	AC011253 Drosophill	890	18	0.6 241886	10	AC027285	AC027285 Mus muscu
c 818	18	0.6 197751	2	AC073290	AC073290 Mus muscu	891	18	0.6 242030	2	AC122673	AC122673 Rattus no
c 819	18	0.6 198226	8	ATCHRIV43	AL161543 Arabidops	c 892	18	0.6 244164	3	AE003548	AE003548 Drosophill
c 820	18	0.6 198348	2	AL772277	AL772277 Mus muscu	c 893	18	0.6 247850	10	AJ421478	AJ421478 Mus muscu
c 821	18	0.6 198453	9	AC093889	AC093889 Homo sapi	894	18	0.6 250991	2	AC127270	AC127270 Mus muscu
c 822	18	0.6 199133	9	AC122212	AC122212 Mus muscu	895	18	0.6 256242	3	AE003756	AE003756 Drosophill
c 823	18	0.6 199280	8	ATCHRIV35	AL161535 Arabidops	896	18	0.6 256373	2	AC020870	AC020870 Mus muscu
c 824	18	0.6 200441	2	AC117150	AC117150 Rattus no	c 897	18	0.6 260050	1	RME603643	AL603643 Rattus no
825	18	0.6 200966	10	AL596108	AL596108 Mouse DNA	c 898	18	0.6 261156	2	AC105829	AC105829 Rattus no
826	18	0.6 201652	2	AC117133	AC117133 Rattus no	c 899	18	0.6 263744	2	AF401201	AF401201 Mus muscu
827	18	0.6 201766	2	AL845260	AL845260 Mus muscu	c 900	18	0.6 263859	2	AC025353	AC025353 Mus muscu
c 828	18	0.6 201989	9	AC073169	AC073169 Homo sapi	c 901	18	0.6 264853	3	AE003644	AE003644 Drosophill
c 829	18	0.6 202685	2	AC127053	AC127053 Rattus no	c 902	18	0.6 269081	2	AC068493	AC068493 Mus muscu
830	18	0.6 203286	10	AL606829	AL606829 Mouse DNA	c 903	18	0.6 271519	2	AC020887	AC020887 Mus muscu
831	18	0.6 203914	2	AC130005	AC130005 Homo sapi	c 904	18	0.6 273785	1	SME591793	AL591793 Sinorhizo
c 832	18	0.6 204317	2	AC117614	AC117614 Mus muscu	c 905	18	0.6 276829	2	AC006741	AC006741 Caenorhab
c 833	18	0.6 204550	2	AL844199	AL844199 Mus muscu	906	18	0.6 278708	3	AE003535	AE003535 Drosophill
c 834	18	0.6 204708	2	AC091617	AC091617 Rattus no	c 907	18	0.6 291622	3	AE006796	AE006796 Caenorhab
835	18	0.6 204775	2	AC120549	AC120549 Mus muscu	c 908	18	0.6 291829	3	AE003525	AE003525 Drosophill
836	18	0.6 205397	2	AC129193	AC129193 Mus muscu	c 909	18	0.6 293431	2	PFMAL13P4	AL049181 Plasmodiu
837	18	0.6 205453	2	AC079556	AC079556 Mus muscu	c 910	18	0.6 296491	3	AE003603	AE003603 Drosophill
838	18	0.6 205577	2	AC011237	AC011237 Homo sapi	c 911	18	0.6 296947	3	AE003421	AE003421 Drosophill
839	18	0.6 205749	10	AL671187	AL671187 Mouse DNA	c 912	18	0.6 299782	2	AC006844	AC006844 Caenorhab
840	18	0.6 206082	2	AC126936	AC126936 Mus muscu	c 913	18	0.6 300962	3	DROSADH04	AE003410 Drosophill
841	18	0.6 206125	10	AL6772219	AL6772219 Mouse DNA	c 914	18	0.6 302797	3	AE003488	AE003488 Drosophill
842	18	0.6 208197	9	AC009107	AC009107 Homo sapi	c 915	18	0.6 306321	3	AE003585	AE003585 Drosophill
843	18	0.6 208327	2	AC119061	AC119061 Bos tauru	916	18	0.6 309023	3	AE003671	AE003671 Drosophill
844	18	0.6 208330	2	AC022327	AC022327 Mus muscu	c 917	18	0.6 310092	3	AE003826	AE003826 Drosophill
c 845	18	0.6 208804	2	AC117025	AC117025 Rattus no	c 918	18	0.6 318200	1	MLRPTNG6	AL583922 Mycobacte
c 846	18	0.6 209164	8	ATFCA5	297340 Arabidopsis	c 919	18	0.6 323450	1	SME591790	AL591790 Sinorhizo
847	18	0.6 209216	2	AC117126	AC117126 Rattus no	c 920	18	0.6 327773	1	AP002554	AP002554 Escherich
848	18	0.6 209600	2	AC022452	AC022452 Mus muscu	921	18	0.6 329100	1	SME591787	AL591787 Sinorhizo
c 849	18	0.6 210336	2	AC022171	AC022171 Homo sapi	922	18	0.6 329861	1	NMA5522491	AL162756 Neisseria
850	18	0.6 210376	2	AL136118	AL136118 Homo sapi	923	18	0.6 349980	6	AX044032	AX044032 Sequence
c 851	18	0.6 210514	10	AC117199	AC117199 Mus muscu	c 924	18	0.6 349980	6	AX344551	AX344551 Sequence
852	18	0.6 211084	2	AC092084	AC092084 Bos tauru	c 925	17	0.5 54	6	AX019625	AX019625 Sequence
c 853	18	0.6 211217	2	AC123840	AC123840 Mus muscu	c 926	17	0.5 54	6	AX019626	AX019626 Sequence
c 854	18	0.6 212085	10	AL663030	AL663030 Mouse DNA	927	17	0.5 77	6	AR104674	AR104674 Sequence
855	18	0.6 213251	2	AC113290	AC113290 Mus muscu	928	17	0.5 77	6	I38257	I38257 Sequence 87
856	18	0.6 213452	2	AL663090	AL663090 Mus muscu	c 929	17	0.5 78	8	TOMTRALTB	L24059 Lycopersico
857	18	0.6 214166	9	AC009660	AC009660 Homo sapi	c 930	17	0.5 106	9	HSCALMT1	U16851 Homo sapien
c 858	18	0.6 214967	2	AC130539	AC130539 Mus muscu	c 931	17	0.5 107	6	AX019620	AX019620 Sequence
c 859	18	0.6 215050	1	AL646057	AL646057 Ralstonia	932	17	0.5 117	6	AR168401	AR168401 Sequence
c 860	18	0.6 216776	10	AL669920	AL669920 Mouse DNA	c 933	17	0.5 117	9	HS299552	299552 Homo sapien
861	18	0.6 217017	2	AC124385	AC124385 Mus muscu	c 934	17	0.5 120	6	AX019627	AX019627 Sequence
862	18	0.6 217130	2	AC123838	AC123838 Mus muscu	c 935	17	0.5 126	6	AX435716	AX435716 Sequence
863	18	0.6 217352	10	AL645784	AL645784 Mouse DNA	c 936	17	0.5 157	9	HSAX01520	AF009460 Homo sapi
864	18	0.6 217420	1	BSUB0017	299120 Bacillus su	c 937	17	0.5 170	5	STU60219	U60219 Salmo trutt
c 865	18	0.6 218193	2	AC121608	AC121608 Mus muscu	c 938	17	0.5 176	6	AX019585	AX019585 Sequence
866	18	0.6 218209	2	AC099425	AC099425 Rattus no	c 939	17	0.5 214	8	AY038211	AY038211 Sinodiels
867	18	0.6 218298	2	AC116389	AC116389 Mus muscu	c 940	17	0.5 216	8	AF073563S1	AF073563 Naufraga
c 868	18	0.6 219709	2	AC091360	AC091360 Rattus no	941	17	0.5 224	8	AY020225	AY020225 Oryza sat





Db 421 TCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAAGACACGGCTATTACGAA 480  
Qy 481 AATTACAAAACATATACGCGTATATGCGGAAGAGCGCCTGAAGACAGGAGCGGTAAA 540  
Db 481 AATTACAAAACATATACGCGTATATGCGGAAGAGCGCCTGAAGACAGGAGCGGTAAA 540  
Qy 541 GACATTAAAGCTTCTTCGACGATGAGGCGGTATAGAGACTGAAGCAAGCCGACGGAT 600  
Db 541 GACATTAAAGCTTCTTCGACGATGAGGCGGTATAGAGACTGAAGCAAGCCGACGGAT 600  
Qy 601 ATCCGCCAGTAAAAGAAATCGACACATCGATGTGCTCCCATATTATTGGCGGGGT 660  
Db 601 ATCCGCCAGTAAAAGAAATCGACACATCGATGTGCTCCCATATTATTGGCGGGGT 660  
Qy 661 TCCGTCGGCAGACCTGACGCGGTATTTGGCCCGGATGCGACGCTACACATAATGAAT 720  
Db 661 TCCGTCGGCAGACCTGACGCGGTATTTGGCCCGGATGCGACGCTACACATAATGAAT 720  
Qy 721 ACGCATGATGGAACCAAGAACGAATAATGCTGAGCCATCGCAATGCAATGCAATGCAAT 780  
Db 721 ACGCATGATGGAACCAAGAACGAATAATGCTGAGCCATCGCAATGCAATGCAATGCAAT 780  
Qy 781 CTGGCGAAGCTGGCGTGGCATCTCAATACAGCTTTTGAACACATCGAGGCGAGC 840  
Db 781 CTGGCGAAGCTGGCGTGGCATCTCAATACAGCTTTTGAACACATCGAGGCGAGC 840  
Qy 841 ACTGCGGACCATTTCCAAATAGCCAAATTCGGAGGAGCAGTACCGCAAGCGTTGCTCGCC 900  
Db 841 ACTGCGGACCATTTCCAAATAGCCAAATTCGGAGGAGCAGTACCGCAAGCGTTGCTCGCC 900  
Qy 901 TATTCCGCGGTGATAAACAGACGAGGATATCCGCCGTATGCAACAGAGCGATTACGCG 960  
Db 901 TATTCCGCGGTGATAAACAGACGAGGATATCCGCCGTATGCAACAGAGCGATTACGCG 960  
Qy 961 AACTTGTCTTACCACATCGTAATAAACAATGCTTTTCAATTTTTCGCAAGCAATGAC 1020  
Db 961 AACTTGTCTTACCACATCGTAATAAACAATGCTTTTCAATTTTTCGCAAGCAATGAC 1020  
Qy 1021 GCACAAGCTCAGCCCAACACACTGACCTATTGCCATTTTATGAAAAAGATGCTCAAAA 1080  
Db 1021 GCACAAGCTCAGCCCAACACACTGACCTATTGCCATTTTATGAAAAAGATGCTCAAAA 1080  
Qy 1081 GGCATTATCAGTGCAGGCGGTAGACCGCAGTGGAGAAAAAGTTCAATGGCTTCCAAACAT 1140  
Db 1081 GGCATTATCAGTGCAGGCGGTAGACCGCAGTGGAGAAAAAGTTCAATGGCTTCCAAACAT 1140  
Qy 1141 TCCGGAATTACTGCTGCTGCTATCGCACCTATGAAGCAAGCGTCCGTTTCACC 1200  
Db 1141 TCCGGAATTACTGCTGCTGCTATCGCACCTATGAAGCAAGCGTCCGTTTCACC 1200  
Qy 1201 CGTACAAACCCGATTCAAAATGCGGAAACATCCTTTTCCGACCCCATCGTAACCGGCAG 1260  
Db 1201 CGTACAAACCCGATTCAAAATGCGGAAACATCCTTTTCCGACCCCATCGTAACCGGCAG 1260  
Qy 1261 GCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAGACGAACTCGGTACCAAGCTG 1320  
Db 1261 GCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAGACGAACTCGGTACCAAGCTG 1320  
Qy 1321 CTGACAAAGCTCAGGACATCGGTGCTGCTGCTGAGCAGCAAGTTTCCGCTGGGGACTG 1380  
Db 1321 CTGACAAAGCTCAGGACATCGGTGCTGCTGCTGAGCAGCAAGTTTCCGCTGGGGACTG 1380  
Qy 1381 CTGGATGCGGTTAAGGCCATGAAGGACCCGCGTCTTTCCGTTCCGCGACTTTTACC GCC 1440  
Db 1381 CTGGATGCGGTTAAGGCCATGAAGGACCCGCGTCTTTCCGTTCCGCGACTTTTACC GCC 1440  
Qy 1441 GATAGAAAGGTACATCCGATATTGCTTACTCTTCCGTAACGACATTTTCAGGCACGGGC 1500  
Db 1441 GATAGAAAGGTACATCCGATATTGCTTACTCTTCCGTAACGACATTTTCAGGCACGGGC 1500  
Qy 1501 GGCCTGATCAAAAAGGCGGACGCACTGCACTGACGCAACACACCTATACGGGC 1560  
Db 1501 GGCCTGATCAAAAAGGCGGACGCACTGCACTGACGCAACACACCTATACGGGC 1560

Qy 1561 AAAACCATTATCGAAGGCGGTTTCGCTGGTGTGTATACGGCAACAACAAATCGGATATGCGC 1620  
Db 1561 AAAACCATTATCGAAGGCGGTTTCGCTGGTGTGTATACGGCAACAACAAATCGGATATGCGC 1620  
Qy 1621 GTCGAAACAAAGGTGCGCTGATTTATACGGGGGGGATCCGGGGTAGCCTGAACAGC 1680  
Db 1621 GTCGAAACAAAGGTGCGCTGATTTATACGGGGGGGATCCGGGGTAGCCTGAACAGC 1680  
Qy 1681 GACGCGATTGCTATCTGGCAGATACCGACCGATCCGGGGGCAAAACGAAACCGTGACATC 1740  
Db 1681 GACGCGATTGCTATCTGGCAGATACCGACCGATCCGGGGGCAAAACGAAACCGTGACATC 1740  
Qy 1741 AAAGTCGATCTGACGCTGGGGCGCAAGGTACGCTGTACACAGTTTGGGCAAACTGCTG 1800  
Db 1741 AAAGTCGATCTGACGCTGGGGCGCAAGGTACGCTGTACACAGTTTGGGCAAACTGCTG 1800  
Qy 1801 AAAGTCGATCTGACGCTGGGGCGCAAGGTACGCTGTACATGTCGGCACGCGGCAAGGG 1860  
Db 1801 AAAGTCGATCTGACGCTGGGGCGCAAGGTACGCTGTACATGTCGGCACGCGGCAAGGG 1860  
Qy 1861 GCAGGCTATCTCAACCGTACCGGACAACTGTTCCCTTCTCTGAGTGCAGTGCAGTCCG 1920  
Db 1861 GCAGGCTATCTCAACCGTACCGGACAACTGTTCCCTTCTCTGAGTGCAGTGCAGTCCG 1920  
Qy 1921 CGGGATTATCTTCTTCCAAAAACATCGAAACCGAGGTGCTGCTGGCTTCCCTCGAC 1980  
Db 1921 CGGGATTATCTTCTTCCAAAAACATCGAAACCGAGGTGCTGCTGGCTTCCCTCGAC 1980  
Qy 1981 AGCGTCGAAAAAAGAGCGGCGAGTGAAGCGACACGCTGCTCTATTATGTCGTCGCGC 2040  
Db 1981 AGCGTCGAAAAAAGAGCGGCGAGTGAAGCGACACGCTGCTCTATTATGTCGTCGCGC 2040  
Qy 2041 AATCGCGCACGAGTCTTCCGACGCGCACATTCGCGCGCGCGCGCTGTAACACAGCC 2100  
Db 2041 AATCGCGCACGAGTCTTCCGACGCGCACATTCGCGCGCGCGCGCTGTAACACAGCC 2100  
Qy 2101 GTAGAAACAGGCGGCGAGCAATCTCGAAAACTGTAGTGGTGAACCTGATGCTCCGAATCA 2160  
Db 2101 GTAGAAACAGGCGGCGAGCAATCTCGAAAACTGTAGTGGTGAACCTGATGCTCCGAATCA 2160  
Qy 2161 TCCGCAACACCCGAGACGGTTGAAACTCGCGCGCGCGCGCACAGATATCGCGGCGATC 2220  
Db 2161 TCCGCAACACCCGAGACGGTTGAAACTCGCGCGCGCGCGCACAGATATCGCGGCGATC 2220  
Qy 2221 GCGCCTACGGGCAACTTTCCGCGCAGCGGCGGTACAGCATGCGAATGCGCGCGAC 2280  
Db 2221 GCGCCTACGGGCAACTTTCCGCGCAGCGGCGGTACAGCATGCGAATGCGCGCGAC 2280  
Qy 2281 GGTGTACGCATCTTCAACAGTCTCCCGCTATCGCGCTATCGGACAGTACGCGCGCCAT 2340  
Db 2281 GGTGTACGCATCTTCAACAGTCTCCCGCTATCGCGCTATCGGACAGTACGCGCGCCAT 2340  
Qy 2341 GCGGATATGAGGAGCGCGGCTGAAACCGGTGAAACCGGTGGAACGTTGGGAAACGCGGTGT 2400  
Db 2341 GCGGATATGAGGAGCGCGGCTGAAACCGGTGAAACCGGTGGAACGTTGGGAAACGCGGTGT 2400  
Qy 2401 GGTCTGCGCGTATCGCGCAACCAACAGGACGTTGAAACGTTGGGAAACGCGGTGT 2460  
Db 2401 GGTCTGCGCGTATCGCGCAACCAACAGGACGTTGAAACGTTGGGAAACGCGGTGT 2460  
Qy 2461 GAAGCAAAATGCGCGCAGTACCCAAACCGTTCGCGCATTTGCCGCAAAACCGCGGCAAAAT 2520  
Db 2461 GAAGCAAAATGCGCGCAGTACCCAAACCGTTCGCGCATTTGCCGCAAAACCGCGGCAAAAT 2520  
Qy 2521 ACGACAGACGCGCCACTCTGGGATGGGACACAGCATGGAGGAAAAACAGTGCAT 2580  
Db 2521 ACGACAGACGCGCCACTCTGGGATGGGACACAGCATGGAGGAAAAACAGTGCAT 2580  
Qy 2581 GCAAAAACCGACGATTAGTCTGTTTCAGGATACGCGCAGGATGCGGGGATATCGGC 2640  
Db 2581 GCAAAAACCGACGATTAGTCTGTTTCAGGATACGCGCAGGATGCGGGGATATCGGC 2640

Qy	2641	TATCTCAAGGCGCTGTTCTCTCTACGGACGCTACAAAAACAGCATCAGCCCGACGACCGGT	2700
Ds	2641	TATCTCAAGGCGCTGTTCTCTCTACGGACGCTACAAAAACAGCATCAGCCCGACGACCGGT	2700
Qy	2701	CGGACGCAACATGCGGAAGGACGCTCAACGGCAGCTGATGACGCTGGCGGCACTGGC	2760
Ds	2701	CGGACGCAACATGCGGAAGGACGCTCAACGGCAGCTGATGACGCTGGCGGCACTGGC	2760
Qy	2761	GGTCTCAACGCTTCGTTTGGCGCAACGGGAGATTGACGGTGAAGCGGTCTCGCGTAC	2820
Ds	2761	GGTCTCAACGCTTCGTTTGGCGCAACGGGAGATTGACGGTGAAGCGGTCTCGCGTAC	2820
Qy	2821	GACCTGCTCAACAGGATGCAATCCGCCCAAAAGGCACTGTTGGCTGGAGCGGCAAC	2880
Ds	2821	GACCTGCTCAACAGGATGCAATCCGCCCAAAAGGCACTGTTGGCTGGAGCGGCAAC	2880
Qy	2881	AGCCTCACTGAAGCAGCAGCTGCTGGGACTCGCGGCTGAAGCTGTGCAACCCCTTGAGC	2940
Ds	2881	AGCCTCACTGAAGCAGCAGCTGCTGGGACTCGCGGCTGAAGCTGTGCAACCCCTTGAGC	2940
Qy	2941	GATAAAGCGCTCCTGTTTGAACAGCGGGCGTGAACCGGACCTGAACGGACGGGACTAC	3000
Ds	2941	GATAAAGCGCTCCTGTTTGAACAGCGGGCGTGAACCGGACCTGAACGGACGGGACTAC	3000
Qy	3001	ACGTAACGGGCGGCTTTACCGGGCGGACTGCAACACCGGCAAGACGGGGGCAACGCAAT	3060
Ds	3001	ACGTAACGGGCGGCTTTACCGGGCGGACTGCAACACCGGCAAGACGGGGGCAACGCAAT	3060
Qy	3061	ATGCCGACACACCCCGCTGTTGCGGCTGTCGGCGGATGTCGAATTCGGCAACGGCTGG	3120
Ds	3061	ATGCCGACACACCCCGCTGTTGCGGCTGTCGGCGGATGTCGAATTCGGCAACGGCTGG	3120
Qy	3121	AACGGCTTGGCACCTTACAGCTACGCGGTTTCCAAACAGTACGGCAACACGCGGACGA	3180
Ds	3121	AACGGCTTGGCACCTTACAGCTACGCGGTTTCCAAACAGTACGGCAACACGCGGACGA	3180
Qy	3181	GTCGGCTAGGCTACCGGTTCTGA	3204
Ds	3181	GTCGGCTAGGCTACCGGTTCTGA	3204
RESULT 2			
NME311654			
LOCUS			
DEFINITION			
Neisseria meningitidis ausp gene for autotransporter serine protease.			
ACCESSION			
AJ311654			
VERSION			
AJ311654.1 GI:13397949			
KEYWORDS			
ausp gene; autotransporter serine protease.			
SOURCE			
Neisseria meningitidis			
ORGANISM			
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.			
REFERENCE			
1			
Turner, D.P., Wooldridge, K.G. and Ala'Aldeen, D.A.			
Autotransported Serine Protease A of Neisseria meningitidis: an			
Immunogenic, Surface-Exposed Outer Membrane, and Secreted Protein			
JOURNAL			
Infect. Immun. 70 (8), 4447-4461 (2002)			
MEDLINE			
22112893			
REFERENCE			
2. (bases 1 to 3204)			
Ala'Aldeen, D.A.			
AUTHORS			
Direct Submission			
TITLE			
Submitted (13-MAR-2001) Ala'Aldeen D.A., Microbiology, University			
JOURNAL			
of Nottingham, University Hospital, Nottingham, NG7 2UH, UNITED			
KINGDOM			
Related sequence: AJ277537.			
COMMENT			
FEATURES			
source			
1. .3204			
Location/Qualifiers			
/organism="Neisseria meningitidis"			
/strain="Group C Et37"			
/isolate="Z4181"			
/db_xref="taxon:487"			
/country="Mali"			

gene	1. .3204	830 a	909 c	900 g	565 t
CDS	/gene="ausp"				
	1. .3204				
	/gene="ausp"				
	/codon_start=1				
	/transl_table=11				
	/product="autotransporter serine protease (Ausp)"				
	/protein_id="CAC34605.1"				
	/db_xref="GI:13397950"				
	/db_xref="SPTREMBL:O9AE78"				
	/translation="MRTTPTFTKTKFPAAMALAVATTLSACLGCGGGTSAPDFNAGG				
	TGIGNSRATTAASVYAGIKNEMKDRMLCAGDDVAVTDRDAKINAPPPNLT				
	GFETNPNDAYKNLILNPALEAGYGRGVEVIGVDTGSEGVSTSPFLYKREHPYNE				
	NKNTAYMRKEAPEDGGDKIKASDDDEAVLETAKPTDIRHVKFIGHIDVVSHTIG				
	GRSVDRPAGGIAPDATHIMTHDGTKEIMSAARNAWVKLGERGVRIVNSFGTT				
	FRAGTDAQFOIANSEQYQALLAYSGGDKTDEGIRLMOOSDYGNLSYHNRKNMLFI				
	SSANLQAPQNTLALLPEYKDAQGIITVAGVDRSGKFNHCHGTAAWMLCSAP				
	YFASVFTNPTIQTAGTSPAPIVTGTAAALLQKYPWMSNDNLRTLLTTLTAQDIGAV				
	GVDSKFGWLLDAGRAMGPSFPFGDFTADTKGSDIAYSFNRDTSGTGGI KKGGS				
	OLQLHGNNTYTGKT IIEGSLVLYGNKSDMBVETKGLIYNGAASGGSINSDGIWYL				
	AOTDRSGANETVHIKGDLOLGEGTLYTRLGKLLKVDGTAMTGKLYMSARKGAGYL				
	NRTGQRPFLSAAIKGRDYSFFTNIDTGGLLASLDSEVETAGSEGDITUSYVRRGNA				
	ARTASAAHSAPAGLKHAVEOGGSNLENLENMELDASESATPETVETAAARDTMDPGI				
	RPYGATFRAAAAYQHANAADGVRIENSLAATVYADSAHADMQRLKAVSDGLDHN				
	ATGLRVIAOTQODGGTWEQGVGKMRGTSOTVIGIAAKTGENTTAATLGMHSTWSE				
	NSANAKTDSISLFIAGTIDHGDIGLYLKGLFSYGRYKNSISRSTGADDEHAGSVNGLM				
	OLGALGVNVFPAATGDLTVBGLRLYDILLKQDAFAEKGKSGALNSGNSLFTGLVGLAG				
	LKLSQLSKAVLIFATAGVERDLNDRDITVTGTFGTATATKTKTARNMPHRLVLAGL				
	GADVEFGNWNGLARYSYASKQYGNHSGRVGVGYRF"				
BASE COUNT	830 a	909 c	900 g	565 t	
ORIGIN					
Query Match	98.4%;	Score 3153;	DB 1;	Length 3204;	
Best Local Similarity	100.0%;	Pred. No. 0;			
Matches 3203;	Conservative	0;	Mismatches	1;	Indels 0; Gaps 0;
Qy	1	ATCGAACGACCCCACTTCCCTACAAAACTTCAAAACGGCTGCATGGCGTTAGCT	60		
Ds	1	ATCGAACGACCCCACTTCCCTACAAAACTTCAAAACGGCTGCATGGCGTTAGCT	60		
Qy	61	GTTGCAACAACTTTCTGCTGCTTACGGCGGCGGCGGCGACTTCTGCGCCGACTTC	120		
Ds	61	GTTGCAACAACTTTCTGCTGCTTACGGCGGCGGCGGCGACTTCTGCGCCGACTTC	120		
Qy	121	AATGCCAGCGCGCCGCTATCGGCAGCAACGAGCAGACACAGCAAGAAATCAGCAGCA	180		
Ds	121	AATGCCAGCGCGCCGCTATCGGCAGCAACGAGCAGACACAGCAAGAAATCAGCAGCA	180		
Qy	181	GTATCTTACGCCGCTATCAAGAACGAAATGTGCAAGACAGAGCATGCTCTGTGCCGGT	240		
Ds	181	GTATCTTACGCCGCTATCAAGAACGAAATGTGCAAGACAGAGCATGCTCTGTGCCGGT	240		
Qy	241	CGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAATGCCCCCCCCGAAATCTG	300		
Ds	241	CGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAATGCCCCCCCCGAAATCTG	300		
Qy	301	CATACCGGACACTTTACAAACCCCAATGAGCCATACAGAATTTGATCAACCTCAACCT	360		
Ds	301	CATACCGGACACTTTACAAACCCCAATGAGCCATACAGAATTTGATCAACCTCAACCT	360		
Qy	361	GCAATTGAACGAGCTATACAGGACGCGGGTAGAGTAGGTATCGTCGATACAGGCGAA	420		
Ds	361	GCAATTGAACGAGCTATACAGGACGCGGGTAGAGTAGGTATCGTCGATACAGGCGAA	420		
Qy	421	TCCGTCGGCAGCATATCTTTCCGAACTGTATGGCAGAAAAAACAACGCGCTATAACGAA	480		
Ds	421	TCCGTCGGCAGCATATCTTTCCGAACTGTATGGCAGAAAAAACAACGCGCTATAACGAA	480		
Qy	481	AATTACAAAAAACHTATACGGCGTATATCGCGAAGAAACGCGCTGAAGACGAGCGGTAAA	540		
Ds	481	AATTACAAAAAACHTATACGGCGTATATCGCGAAGAAACGCGCTGAAGACGAGCGGTAAA	540		

QY 541 GACATTAAAGCTTCTTTTCAGACGATGAGGCGGTTATAGAGACTGAAGCAAAAGCCGCGAT 600  
DB 541 GACATTTAAAGCTTCTTTTCAGACGATGAGGCGGTTATAGAGACTGAAGCAAAAGCCGCGAT 600  
QY 601 ATCCGCCACGTAAACAAATCGGACACATCGATGTGGTCTCCCATATTATTGGCGGGCGT 660  
DB 601 ATCCGCCACGTAAAGAAATCGGACACATCGATGTGGTCTCCCATATTATTGGCGGGCGT 660  
QY 661 TCCGTGGAGCGCAGACCTCGAGCGGTATTGCGCCCGATGCGACGCTACACATAATGAAT 720  
DB 661 TCCGTGGAGCGCAGACCTCGAGCGGTATTGCGCCCGATGCGACGCTACACATAATGAAT 720  
QY 721 AGCATGATGGAACCAAGAACGAATAATGCTGCGAGCCATCGCAATGATGGGTCAAG 780  
DB 721 AGCATGATGGAACCAAGAACGAATAATGCTGCGAGCCATCGCAATGATGGGTCAAG 780  
QY 781 CTGGGGGAAGCTGGCGTGGCATCGTCAATACAGTTTGGGAACAACATCGAGGGCAGGC 840  
DB 781 CTGGGGGAAGCTGGCGTGGCATCGTCAATACAGTTTGGGAACAACATCGAGGGCAGGC 840  
QY 841 ACTGCCGACCATTTCCAAATAGCAATTCGGAGGACGATACCGCCAAGCGTTGCTCGCC 900  
DB 841 ACTGCCGACCATTTCCAAATAGCAATTCGGAGGACGATACCGCCAAGCGTTGCTCGCC 900  
QY 901 TATTCGGCGGTGATAAAACAGACGAGGGTATCCGCCCTGATGCAACAGACGATACGGC 960  
DB 901 TATTCGGCGGTGATAAAACAGACGAGGGTATCCGCCCTGATGCAACAGACGATACGGC 960  
QY 961 AACTTGTCTACCATCCGTATATAAACATGCTTTTCATTTTTCGGCAAGCAATGAC 1020  
DB 961 AACTTGTCTACCATCCGTATATAAACATGCTTTTCATTTTTCGGCAAGCAATGAC 1020  
QY 1021 GCACAAAGCTCAGCCCAACACATGACCCCTATTGCCATTTTATGAAAAGATGCTCAAAA 1080  
DB 1021 GCACAAAGCTCAGCCCAACACATGACCCCTATTGCCATTTTATGAAAAGATGCTCAAAA 1080  
QY 1081 GGCATTTATCACAGTCGACGGCGTAGACCGCAGTGGAGAAAAGTTCAATGGCTCCAACCAT 1140  
DB 1081 GGCATTTATCACAGTCGACGGCGTAGACCGCAGTGGAGAAAAGTTCAATGGCTCCAACCAT 1140  
QY 1141 TCGGGAATTTACTGCCATGTGGTGCCTATCGGCACCTATGAAGCAAGCGTCCGTTTCACC 1200  
DB 1141 TCGGGAATTTACTGCCATGTGGTGCCTATCGGCACCTATGAAGCAAGCGTCCGTTTCACC 1200  
QY 1201 CGTACAAACCGGATTCAAATTCGCGGAACATCCTTTTCGCGACCCATCGTAAACCGGACG 1260  
DB 1201 CGTACAAACCGGATTCAAATTCGCGGAACATCCTTTTCGCGACCCATCGTAAACCGGACG 1260  
QY 1261 GCGGCTCTGCTGTCAGAAAATACCGGTGGATGAGCAACGACACCTGCGTACCGCGTGG 1320  
DB 1261 GCGGCTCTGCTGTCAGAAAATACCGGTGGATGAGCAACGACACCTGCGTACCGCGTGG 1320  
QY 1321 CTGACAACGGCTCAGGACATCGTGAGTCGGCGTGGAGCAAGTTCGGTGGGGACTG 1380  
DB 1321 CTGACAACGGCTCAGGACATCGTGAGTCGGCGTGGAGCAAGTTCGGTGGGGACTG 1380  
QY 1381 CTGGATTCGGGGTAAGGCCATGAACGGACCGCGTCCCTTTCCGTTTCGGCGACTTTACCGCC 1440  
DB 1381 CTGGATTCGGGGTAAGGCCATGAACGGACCGCGTCCCTTTCCGTTTCGGCGACTTTACCGCC 1440  
QY 1441 GATACGAAAGGTACATCCGATATTGCGTACTGCTCCGTTAAGCAATTTTCAGCGACGGGC 1500  
DB 1441 GATACGAAAGGTACATCCGATATTGCGTACTGCTCCGTTAAGCAATTTTCAGCGACGGGC 1500  
QY 1501 GGCCTGATCAAAAAGGCGGACCCAACTGCACTGCGCGGCAACCAACCTATACGGGC 1560  
DB 1501 GGCCTGATCAAAAAGGCGGACCCAACTGCACTGCGCGGCAACCAACCTATACGGGC 1560  
QY 1561 AAAACCATTTATCGAAGCGGTTGCTGGTGTGTAGCGCAACAAACAAATTCGGATATGCCG 1620  
DB 1561 AAAACCATTTATCGAAGCGGTTGCTGGTGTGTAGCGGCAACAAACAAATTCGGATATGCCG 1620  
QY 1621 GTCGAAACCAAGGTGCGCTGATTTATAACGGGGGGGCATCCGGGGTAGCCTGAACAGC 1680

DB 1621 GTCEAAACAAAGGTGCGTGTATTAACGGGGGCGCATCCCGGTAGCCTGAACAGC 1680  
QY 1681 GACGGCATTTGTCTATCTGGCAGATACCGACCGATCGGGCAACAAACCGGTGCACATC 1740  
DB 1681 GACGGCATTTGTCTATCTGGCAGATACCGACCGATCGGGCAACAAACCGGTGCACATC 1740  
QY 1741 AAAGGCGATCTGCAGCTGGCGGCGAAGGTACGCTGTACACAGCTTTGGGCAAACTGCTG 1800  
DB 1741 AAAGGCGATCTGCAGCTGGCGGCGAAGGTACGCTGTACACAGCTTTGGGCAAACTGCTG 1800  
QY 1801 AAAGTGGACGGTACGCGGATACCGGCGCAAGCTGTACATGTCCGACGCGCAAGGG 1860  
DB 1801 AAAGTGGACGGTACGCGGATACCGGCGCAAGCTGTACATGTCCGACGCGCAAGGG 1860  
QY 1861 GCAGGCTATCTCAACCGTACCGGACACGCTGTTCCCTTCTTGAGTCCGCGCAAACTCGG 1920  
DB 1861 GCAGGCTATCTCAACCGTACCGGACACGCTGTTCCCTTCTTGAGTCCGCGCAAACTCGG 1920  
QY 1921 CGGATTTATTTCTTCTTCAAAACATCGAAACCGACGGTGTCTGCTGGCTTCCCTCGAC 1980  
DB 1921 CGGATTTATTTCTTCTTCAAAACATCGAAACCGACGGTGTCTGCTGGCTTCCCTCGAC 1980  
QY 1981 AGCGTCAAAAACACAGCGGCGAGTGAAGGCGACACGCTGTCTTATTATGTCTCCGCGC 2040  
DB 1981 AGCGTCAAAAACACAGCGGCGAGTGAAGGCGACACGCTGTCTTATTATGTCTCCGCGC 2040  
QY 2041 AATGCCGCGACGACTGCTTTCGGCAGCGGACATTTCCGCGCGCCCGCTGTGAAACACGCG 2100  
DB 2041 AATGCCGCGACGACTGCTTTCGGCAGCGGACATTTCCGCGCGCCCGCTGTGAAACACGCG 2100  
QY 2101 GTAGAACAGGGCGGACAACTCTGAAACCTGATGGTTCGAACTGGATGCTCCGGAATCA 2160  
DB 2101 GTAGAACAGGGCGGACAACTCTGAAACCTGATGGTTCGAACTGGATGCTCCGGAATCA 2160  
QY 2161 TCCGCAACACCCGAGACGGTTGAAACTTGGCGCGCCGACCGCACAGATATGCCGGGCATC 2220  
DB 2161 TCCGCAACACCCGAGACGGTTGAAACTTGGCGCGCCGACCGCACAGATATGCCGGGCATC 2220  
QY 2221 CGCCCCACGCGGCACTTTCCGCGCAGCGGACGCTACAGCATCGAATGCGCGCGAC 2280  
DB 2221 CGCCCCACGCGGCACTTTCCGCGCAGCGGACGCTACAGCATCGAATGCGCGCGAC 2280  
QY 2281 GGTGTACGCTATCAACAGTCTCGCGCTTACCGTCTATGCCCGACAGTACCGCGGCCAT 2340  
DB 2281 GGTGTACGCTATCAACAGTCTCGCGCTTACCGTCTATGCCCGACAGTACCGCGGCCAT 2340  
QY 2341 GCGGATATCAGGGACGCGGCTGAAAGCGGTATCGACGGGTGGACCAACACGCTACG 2400  
DB 2341 GCGGATATCAGGGACGCGGCTGAAAGCGGTATCGACGGGTGGACCAACACGCTACG 2400  
QY 2401 GGTCTCGCGTCAATCGGCAACCCCAACAGGAGCGGTGGAACTGGGAACAGGGCGGTGT 2460  
DB 2401 GGTCTCGCGTCAATCGGCAACCCCAACAGGAGCGGTGGAACTGGGAACAGGGCGGTGT 2460  
QY 2461 GAAGGCAAAATCGCGGCACTACCCAAACCTCGGCTTGCAGGCAAAACCGCGGCAAAAT 2520  
DB 2461 GAAGGCAAAATCGCGGCACTACCCAAACCTCGGCTTGCAGGCAAAACCGCGGCAAAAT 2520  
QY 2521 AGCAGACGCGCCACACTGGGCACTGGGACACAGACATGGAGCGGAAACAGTGCAGAT 2580  
DB 2521 AGCAGACGCGCCACACTGGGCACTGGGACACAGACATGGAGCGGAAACAGTGCAGAT 2580  
QY 2581 GCAAAAACGACAGCATTTAGTCTGTTTGCAGGCACTACGCGCATCGGGCGGATTCGGC 2640  
DB 2581 GCAAAAACGACAGCATTTAGTCTGTTTGCAGGCACTACGCGCATCGGGCGGATTCGGC 2640  
QY 2641 TATCTCAAGGGCTGTTCTCTACGCGCTACAAAACAGCATACGCGCAGCAGCGGT 2700  
DB 2641 TATCTCAAGGGCTGTTCTCTACGCGCTACAAAACAGCATACGCGCAGCAGCGGT 2700  
QY 2701 GCGGACGAACATCGGGAAGGCAAGCGTCAACGCGCACGCTGATGCACTGGGCGCACTGGC 2760

Db	2701	GGGGACGACATGCGGAAGGACGCCTCAACGGCAGCGCTGATGTCACGCTGGCGCACCTGGCG	2760
Oy	2761	GGTGTCAACGTTCCGTTTCCCGCAACGGGAGATTTCAGCGGTGCGAAGGGGGTCTGCGCTAC	2820
Db	2761	GGTGTCAACGTTCCGTTTCCCGCAACGGGAGATTTCAGCGGTGCGAAGGGGGTCTGCGCTAC	2820
Oy	2821	GACCTGCTCAACAGGATGCTATTCGCCGAAAGGCAAGTGTTCGGCTGGAGCGGCAAC	2880
Db	2821	GACCTGCTCAACAGGATGCTATTCGCCGAAAGGCAAGTGTTCGGCTGGAGCGGCAAC	2880
Oy	2881	AGCCTCACTGAAGGCACACTGGTTCGGACTCGCGGCTCTGAGCTCTCGCAACCCCTTGAGC	2940
Db	2881	AGCCTCACTGAAGGCACACTGGTTCGGACTCGCGGCTCTGAGCTCTCGCAACCCCTTGAGC	2940
Oy	2941	GATAAAGCGCTCTCTTTGCAACGGCGGCGCTGGAAACGCGACTCAACGGACGCGACTAC	3000
Db	2941	GATAAAGCGCTCTCTTTGCAACGGCGGCGCTGGAAACGCGACTCAACGGACGCGACTAC	3000
Oy	3001	ACGGTAACGGGGGGTTCACCGGCGGACTGTCAGCAACCGGCAAGACGGGGGACGCGCAAT	3060
Db	3001	ACGGTAACGGGGGGTTCACCGGCGGACTGTCAGCAACCGGCAAGACGGGGGACGCGCAAT	3060
Oy	3061	ATGCCGACACCCGCTGGTTCGCGGCTGCGCGCGGATGCGAATTCGGCAACGCGCTGG	3120
Db	3061	ATGCCGACACCCGCTGGTTCGCGGCTGCGCGCGGATGCGAATTCGGCAACGCGCTGG	3120
Oy	3121	AACGGCTTGGCAGCTTACAGCTACGCCGCTTCCAAACAGTACGGCAACACACAGCGGACGA	3180
Db	3121	AACGGCTTGGCAGCTTACAGCTACGCCGCTTCCAAACAGTACGGCAACACACAGCGGACGA	3180
Oy	3181	GTGCGGTAGGCTACCGGTTCTGA 3204	
Db	3181	GTGCGGTAGGCTACCGGTTCTGA 3204	
RESULT 3			
NMA222491/c			
LOCUS	349061 bp DNA linear BCT 04-DEC-2000		
DEFINITION	Neisseria meningitidis serogroup A strain 22491 complete genome;		
segment 2/7.			
ACCESSION	AL162753 AL157959		
VERSION	AL162753.2 GI:7379120		
KEYWORDS	Neisseria meningitidis 22491.		
SOURCE	Neisseria meningitidis 22491		
ORGANISM	Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;		
REFERENCE	1 (bases 1 to 349061)		
AUTHORS	Parkhill,J., Achtman,M., James,K.D., Bentley,S.D., Churcher,C., Klee,S.R., Morelli,G., Basham,D., Brown,D., Chillingworth,T., Davies,R.M., Davis,P., Devlin,K., Feltwell,T., Hamlin,N., Holroyd,S., Jagels,K., Leather,S., Moule,S., Mungall,K., Quail,M.A., Rajandream,M.A., Rutherford,K.M., Simmonds,M., Skelton,J., Whitehead,S., Spratt,B.G. and Barrell,B.G.		
TITLE	Complete DNA sequence of a serogroup A strain of Neisseria meningitidis 22491		
JOURNAL	Nature 404 (6777), 502-506 (2000)		
MEDLINE	20222556		
PUBMED	10761919		
REFERENCE	2 (bases 1 to 349061)		
AUTHORS	Parkhill,J.		
TITLE	Direct Submission		
JOURNAL	Submitted (30-MAR-2000) Submitted on behalf of the Neisseria sequencing team, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk		
COMMENT	Notes: Details of N. meningitidis sequencing at the Sanger Centre are available on the World Wide Web. (URL, http://www.sanger.ac.uk/Projects/N_meningitidis/).		
FEATURES	Location/Qualifiers		
source	1..349061		
	/organism="Neisseria meningitidis 22491"		
	/strain="22491"		

/db_xref="taxon:122587"	
/note="serogroup: A"	
complement(24..206)	
/note="ATR repeat; hmms hit to HMM ATR (1 - 183), score: 310.39"	
/label=ATR	
209..212	
gene	
220..681	
/gene="NMA0368"	
220..681	
/gene="NMA0368"	
/note="NMA0368, probable integral membrane protein, len: 153 aa; contains four probable transmembrane domains"	
/codon_start=1	
/transl_table=11	
/product="putative integral membrane protein"	
/protein_id="CAB83669.1"	
/db_xref="GI:7379121"	
/db_xref="SPTREMBL:Q9JWH7"	
/translation="MQQNKRKSPPIVMLLVSVLWIASLSNVAFYLGNGHSMRGLTV LILGSIPLSLDIRYCAVYANVWLAIVLLALRKVKVPVHAAFWGLALVAFSVKAVTV DEAGTSDIVRYGAGFYLWYAAFAVASIGTFAGNKKERKAASADGTKNDV"	
671..1492	
/gene="hemk"	
671..1492	
/gene="hemk"	
/note="NMA0369, hemK, HemK protein, len: 273 aa; similar to e.g. HEMK_ECOLI P37186 HEMK protein (277 aa), fasta scores; E(): 0, 42.3% identity in 279 aa overlap. Contains PS00092 N-6 Adenine-specific DNA methylases signature"	
/codon_start=1	
/transl_table=11	
/product="HemK protein"	
/protein_id="CAB83670.1"	
/db_xref="GI:7379122"	
/db_xref="SPTREMBL:Q9JWH6"	
/translation="MTEDKWLGLSKPKNEARMLLOYVSEYTRVOLITRGEEMPDEV GRVMDRLAQLNGEPVAYILGAREFYGRFTVPNPVSLIPRTEHLEVAFLAPEN RQWDRLGTGCGAVATVALEPDAFVRASDISPPALETKARNAADLGAFAVGSGWF DTDMPSEKWDIIVSNPPYIENGDKHLSQGLRPEPOIALTDFSDGLSCLRTLAOGAP DRLAEGFLLLEHGFDOGAARVGLAENGSGVETLPDLAGLDRTVLGTYMKHLK"	
831..840	
/gene="hemk"	
/note="Core DNA uptake sequence: gccgtctctgaa"	
/label=DUS	
1162..1171	
/gene="hemk"	
/note="Core DNA uptake sequence: gccgtctctgaa"	
/label=DUS	
1187..1207	
/gene="hemk"	
/note="PS00092 N-6 Adenine-specific DNA methylases signature"	
complement(1279..1288)	
/note="Core DNA uptake sequence: gccgtctctgaa"	
/label=DUS	
1565..1568	
gene	
1578..2966	
/gene="NMA0370"	
1578..2966	
/gene="NMA0370"	
/note="NMA0370, probable integral membrane protein, len: 462 aa; similar to hypothetical proteins e.g. Y325_HAEIN P46440 hypothetical protein HT0325 (450 aa), fasta scores; E(): 0, 49.8% identity in 464 aa overlap"	
/codon_start=1	
/transl_table=11	
/product="putative integral membrane protein"	
/protein_id="CAB83671.1"	
/db_xref="GI:7379123"	
/db_xref="SPTREMBL:Q9JWH5"	
/translation="MNAVVAIVMLVLSLRVHVLSLTIGAFVGGVAGAMPQNTIA DAAGQVSOAGIIPVFENKLEGGKIALSYAMLGAFAITAITHSGLPQOLAGAVVRKLNR	

GMPPDSVRSEGVVKKLLLSIIIVMGMMSQNIPIIHIAFIPTMTVPVPLLLLVFNRLKIDR
RLIACVITGLVTTMFLPYGFGAIFELNELLGNHISAAFPQDVKKNINWMAAIAFAL
GMLAGLLAFVHYRKPRLYQSNADTGNADANRPQPSAYRSLAAVAIAVAFATQAL
MYEDSVLIGAMLGFAYRMLGVINRDKANDVFGEGLKMAWVGFIMIAOGFAFVWNA
TGHIOPLVSSMAIFGNSKGMALMLVLLVTMGIGSSFSFPLIITAIYVPLCVGL
GPSPLATVAIVGTAGALGDAGSPASDSTLGTPTMGLNADQHDHRSVIPTFIHYNIP
LLIAGWIRAMVL"
2967. .3191
/genes="slyx"
/notes="slyx"
Similar to SLX protein homolog, len: 74 aa;
fasta scores: E(): 0.0017, 33.8% identity in 74 aa
overlap, and SLX\_ECOLI P30857 SLX protein (72 aa), fasta
scores: E(): 0.24, 32.4% identity in 68 aa overlap"
/codon\_start=1
/transl\_table=11
/product="SLX protein homolog"
/protein\_id="CAB83672.1"
/db\_xref="GI:7379124"
/db\_xref="SPTREMBL:Q9JWH4"
/translation="MDAVOEFHRITELEIQSALQEDVIAGLNAMVAELRQTLDLQQA
QLRLYQKQDRNPDAQEYSLRDEIPPHY"
complement(3271. .3423)
/genes="NMA0372"
complement(3271. .3423)
/notes="NMA0372, unknown, questionable CDS, len: 50 aa"
/codon\_start=1
/transl\_table=11
/product="very hypothetical protein NMA0372"
/protein\_id="CAB83673.1"
/db\_xref="GI:7379125"
/db\_xref="SPTREMBL:Q9JWH3"
/translation="MFGRLSPUGSDGFLLAVYEPVGFICLOGKTAIFYRSGGKRFRHRI
KGILSD"
3562. .3571
/notes="Core DNA uptake sequence: gccgtctctgaa"
/label=DUS
complement(3586. .4356)
/genes="thif"
complement(3586. .4356)
/genes="thif"
/notes="NMA0373, thif, probable Thif protein, len: 256 aa;
similar to e.g. Thif\_ECOLI P30138 Thif protein (251 aa),
fasta scores: E(): 0.431% identity in 246 aa overlap,
and MOEB\_ECOLI P12282 molybdopterin biosynthesis MOEB
protein. (249 aa), fasta scores: E(): 0.43.9% identity in
244 aa overlap (note that N.m. does not have orthologs of
any other molybdopterin biosynthesis proteins). Contains
Pfam match to entry PF00899 Thif\_family, Thif family"
/codon\_start=1
/transl\_table=11
/product="Thif protein"
/protein\_id="CAB83674.1"
/db\_xref="GI:7379126"
/db\_xref="SPTREMBL:Q9JWH2"
/translation="MTTTEHDNDDAFLRLYRSHLLDEIGEQOKLSAHLILVCGG
GLGAALPYLASGIGTLTIADSDTVELNLRQVAFDEGDVGKLTALRLRHLN
HIVDRTINEKLDGRLTLVQNAADIVLDCDNYATRAQVNRACVAKPTPLVSGAAR
FEGQLAYRPDLPDPCYACLFDGSGADGICSLFGVFPPLVGIIGSTQAAEALKILL
DAGEPSHRLAVYRALEGQWQVFDLPNRPNCEPCVGAER"
3787. .3796
/notes="Core DNA uptake sequence: gccgtctctgaa"
/label=DUS
complement(3847. .4257)
/genes="thif"
/notes="Pfam match to entry PF00899 Thif\_family, Thif
family, score 186.60, E-value 4.1e-52"
complement(4418. .4427)
/notes="Core DNA uptake sequence: gccgtctctgaa"
/label=DUS

gene 4471. .7224
CDS /gene="ppc"
4471. .7224
/genes="ppc"
/EC\_number="4.1.1.31"
/notes="NMA0374, ppc, phosphoenolpyruvate carboxylase, len:
917 aa; similar to many e.g. CAPP\_RHOPA O32483
phosphoenolpyruvate carboxylase (pc 4.1.1.31) (936 aa),
fasta scores: E(): 0.43.3% identity in 928 aa overlap.
Contains 2x pfam match to entry PF00311 PEPCase,
Phosphoenolpyruvate carboxylase, PS00017 ATP/Grp-binding
site motif A (P-loop), and PS00393 Phosphoenolpyruvate
carboxylase active site 2"
/codon\_start=1
/transl\_table=11
/product="phosphoenolpyruvate carboxylase"
/protein\_id="CAB83675.1"
/db\_xref="GI:7379127"
Query Match 67.3% Score 2156; DB 1; Length 349061;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 3006; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
Qy 182 TATCTTACGCGGTATCAAGAACGAAATGTGCAAGACAGAGCATGCTCTGTGCGCGTC 241
Db 126136 TATCTTACGCGGTATCAAGAACGAAATGTGCAAGACAGAGCATGCTCTGTGCGCGTC 126077
Qy 242 GGGATGACGTTGCGGTACACACAGGAGTGCACAAATCAATGCCCCCCCGGATCTGC 301
Db 126076 GGGATGACGTTGCGGTACACACAGGAGTGCACAAATCAATGCCCCCCCGGATCTGC 126017
Qy 302 ATACCGGAGACTTTACAAACCCAAATGACGCATACAGAAATTTGATCAACCTCAACCTG 361
Db 126016 ATACCGGAGACTTTACAAACCCAAATGACGCATACAGAAATTTGATCAACCTCAACCTG 125957
Qy 362 CAATTGAAGAGCTATACAGGACGCGGGTAGAGTAGTATCGTCGATACAGCGGAAT 421
Db 125956 CAATTGAAGAGCTATACAGGACGCGGGTAGAGTAGTATCGTCGACACAGCGGAAT 125897
Qy 422 CCGTCGCGCAGCATATCCTTTCCCGAACTGTATGCGAGAAAAGACACGGCTATACGAAA 481
Db 125896 CCGTCGCGCAGCATATCCTTTCCCGAACTGTATGCGAGAAAAGACACGGCTATACGAAA 125837
Qy 482 ATTACAAAACTATACGCGGTATATGCGAAGAGAGCGCTGAAGACGAGCGGTAAG 541
Db 125836 ATTACAAAACTATACGCGGTATATGCGAAGAGAGCGCTGAAGACGAGCGGTAAG 125777
Qy 542 ACATTAAGCTTCTTCGACATGAGCGCTTATAGAGACTGAAGCAAGCCGCGGATA 601
Db 125776 ACATTAAGCTTCTTCGACATGAGCGCTTATAGAGACTGAAGCAAGCCGCGGATA 125717
Qy 602 TCCGCCACGTAAAGAAATCGACACATCGATGTGGTCTCCCATATTTATTCGCGCGCTT 661
Db 125716 TCCGCCACGTAAAGAAATCGACACATCGATGTGGTCTCCCATATTTATTCGCGCGCTT 125657
Qy 662 CCGTGGACGCGCAGACCTGCAGCGGTATTTGCGCCCGATGCGACGCTACACATAATGAATA 721
Db 125656 CCGTGGACGCGCAGACCTGCAGCGGTATTTGCGCCCGATGCGACGCTACACATAATGAATA 125597
Qy 722 CGCATGATGAACCAAGAACGAAATATGCTCGAGCCATCCGCAATGCGGTGCAAGC 781
Db 125596 CGCATGATGAACCAAGAACGAAATATGCTCGAGCCATCCGCAATGCGGTGCAAGC 125537
Qy 782 TGGCGCAACGTGGCGTGGCATCTCAATACAGTCTTTTGGAAACAACATCGAGGCGAGCA 841
Db 125536 TGGCGCAACGTGGCGTGGCATCTCAATACAGTCTTTTGGAAACAACATCGAGGCGAGCA 125477
Qy 842 CTGCGGACCATTTCCAAATAGCCAAATTCGGAGGAGCAGTACCGCAAGCGTTGCTCGCT 901
Db 125476 CTGCGGACCATTTCCAAATAGCCAAATTCGGAGGAGCAGTACCGCAAGCGTTGCTCGCT 125417
Qy 902 ATTCGCGCGGTATGATAAACAAGACAGGAGGTATCCGCTGATGCAACAAGAGCGATTACGCA 961

Db 125416 ATTCGGCGGTGATAAAACAGACGAGGGTATCCGCTGATGCAACAGACGATACGGCA 125357  
QY 962 ACTTGTCCTACACATCCGCTAATAAAAAATGCTTTTTCATTTTTCGGAAGCAATGACG 1021  
Db 125356 ACTTGTCCTACACATCCGCTAATAAAAAATGCTTTTTCATTTTTCGGAAGCAATGACG 125297  
QY 1022 CACAAGCTCAGCCCAACACACTGACCTATATGCCATTTTATGAAAAGATGCTCAAAAAG 1081  
Db 125296 CACAAGCTCAGCCCAACACACTGACCTATATGCCATTTTATGAAAAGAGCGCTCAAAAAG 125237  
QY 1082 GCATTATACAGTCGACGCGTAGACCGAGTGGAGAAAAGTTCAATGGCTCCAAACCAT 1141  
Db 125236 GCATTATACAGTTGACGCGTAGACCGAGTGGAGAAAAGTTCAATGGCTCCAAACCAT 125177  
QY 1142 GCGGAATTTACTGCGCATGTGCTGCTATCGCACCCCTATGAAGCAAGCTCCGTTTCACCC 1201  
Db 125176 GCGGAATTTACTGCGCATGTGCTGCTGCTGCGCACCCCTATGAAGCAAGCTCCGTTTCACCC 125117  
QY 1202 GTACAAACCCGATTCAAATTTGCCGGAACATTCCTTTTCCGCACCCATCGTAACCGGCACGG 1261  
Db 125116 GTACAAACCCGATTCAAATTTGCCGGAACATTCCTTTTCCGCACCCATCGTAACCGGCACGG 125057  
QY 1262 CGGCTCTGCTGCGAAGAAATACCCGTGATGAGCAACGACACCTGCGTACCACGCTGC 1321  
Db 125056 CGGCTCTGCTGCGAAGAAATACCCGTGATGAGCAACGACACCTGCGTACCACGCTGC 124997  
QY 1322 TGACAACGGCTCAGGACATCGGTGTCAGTCCGGCTGGACAGCAAGTTTCGGCTGGGAGCTGC 1381  
Db 124996 TGACAACGGCTCAGGACATCGGTGTCAGTCCGGCTGGACAGCAAGTTTCGGCTGGGAGCTGC 124937  
QY 1382 TGGATCGGGTAAAGGCCATGAACGGACCCGCGTCTTTCCGTTTCGGCGACTTTTACCGCGCG 1441  
Db 124936 TGGATACGGGTAAAGGCCATGAACGGACCCGCGTCTTTCCGTTTCGGCGACTTTTACCGCGCG 124877  
QY 1442 ATACGAAAGGTACATCGGATATGCGTACTCTCTCGTACGACATTTTACGGCACCGGGCG 1501  
Db 124876 ATACGAAAGGTACATCGGATATGCGTACTCTCTCGTACGACATTTTACGGCACCGGGCG 124817  
QY 1502 GCCTGATCAAAAAGCGGCGAGCAACTGCACTGCACGGCAACAACACCTATACGGGCA 1561  
Db 124816 GCCTGATCAAAAAGCGGCGAGCAACTGCACTGCACGGCAACAACACCTATACGGGCA 124757  
QY 1562 AAACCATTTATCGAAGCGGTTGCTGGTGTGTTACGGCAACAACAATCGGATATCGCG 1621  
Db 124756 AAACCGTTTATCGAAGCGGTTGCTGGTGTGTTACGGCAACAACAATCGGATATCGACG 124697  
QY 1622 TCGAAACCAAGGTGCGCTGATTTATAAGGGCGGCGATCCGCGGTAGCTGTAACAGCG 1681  
Db 124696 TCGAAACCAAGGTGCGCTGATTTATAAGGGCGGCGATCCGCGGTAGCTGTAACAGCG 124637  
QY 1682 ACGGCATTTGCTATCTGGCAGATACCGACCGATCCGGCGCAACACCGTGCACATCA 1741  
Db 124636 ACGGCATTTGCTATCTGGCAGATACCGACCGATCCGGCGCAACACCGTGCACATCA 124577  
QY 1742 AAGGCGATCTGCAAGCTGGCGGCGGAAGGTACGCTGTACACACGTTTGGGCAAACTGCTGA 1801  
Db 124576 AAGGCGATCTGCAAGCTGGCGGCGGAAGGTACGCTGTACACACGTTTGGGCAAACTGCTGA 124517  
QY 1802 AAGTGACCGGTACGGCGATGACGGCGGCGCAAGCTGTACATGTCCGACACGGGCAAGGGG 1861  
Db 124516 AAGTGACCGGTACGGCGATGACGGCGGCGCAAGCTGTACATGTCCGACACGGGTAAGGGG 124457  
QY 1862 CAGGCTATCTCAACCGTACCGGACAAGTGTTCCTCTCTGAGTCCGCCCAAAATCGGGC 1921  
Db 124456 CAGGCTATCTCAACCGTACCGGACAAGTGTTCCTCTCTGAGTCCGCCCAAAATCGGGC 124397  
QY 1922 GGGATTTATTTCTTTCACAAACATCGAAACCGACGGTGTCTGCTGGCTTCCCTCGACA 1981  
Db 124396 GGGATTTATTTCTTTCACAAACATCGAAACCGACGGGTCGTCTGCTTCCCTCGACA 124337  
QY 1982 GCGTCAAAAACAGCGGGCAGTGAAGGGCAGACGCTGTCTTATTTATGCTCCGTCGGCGCA 2041  
Db 124336 GCGTCAAAAACAGCGGGTAGTGAAGGGCAGACGCTGTCTTATTTATGCTCCGTCGGCGCA 124277

QY 2042 ATCGGCAACGGACTGCTTGGCAGCGGCACATTCGCGCCCGCGCTCTGAAAACAGCGCG 2101  
Db 124276 ATCGGCAACGGACTGCTTGGCAGCGGCACATTCGCGCCCGCGCTCTGAAAACAGCGCG 124217  
QY 2102 TAGAACAGGGCGGCAGCAATCTTGAAAAAACCCTGATGTGCAACTGGATGCTCCGAATCAT 2161  
Db 124216 TAGAACAGGGCGGCAGCAATCTTGAAAAAACCCTGATGTGCAACTGGATGCTCCGAATCAT 124157  
QY 2162 CCGCAACACCCGAGAGCGTTGAAACTGCGGCCCGCCAGCCACAGATATGCCGGGATCC 2221  
Db 124156 CCGCAACACCCGAGAGCGTTGAAACTGCGGCCCGCCAGCCACAGATATGCCGGGATCC 124097  
QY 2222 GCCCCTACCGCGCAACTTTTCGCGCAGCGCAGCCGTACAGCATGCGAATGCCGCCGACG 2281  
Db 124096 GCCCCTACCGCGCAACTTTTCGCGCAGCGCAGCCGTACAGCATGCGAATGCCGCCGACG 124037  
QY 2282 GTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACGCCGCCCATG 2341  
Db 124036 GTGTACGCATCTTCAACAAATCTCGCCGCTACCGTCTATGCCGACAGTACCGGCCCATG 123977  
QY 2342 CCGATATGACGGACCGCCGGCTGAAAAGCCGTATCGGACGGGTTGGACCAACAGCTACGG 2401  
Db 123976 CCGATATGACGGACCGCCGGCTGAAAAGCCGTATCGGACGGGTTGGACCAACAGCTACGG 123917  
QY 2402 GTCTGCGCGTCAATCGCGCAAAACCAACAGCAGCGGTGGAACAGTGGGAACAGGGCGGTGTTG 2461  
Db 123916 GTCTGCGCGTCAATCGCGCAAAACCAACAGCAGCGGTGGAACAGTGGGAACAGGGCGGTGTTG 123857  
QY 2462 AAGGCAAAATGCGCGCAGTACCCAAACCGTCGGCATTCGCCGGAACACCGGGCAAAAATA 2521  
Db 123856 AAGGCAAAATGCGCGCAGTACCCAAACCGTCGGCATTCGCCGGAACACCGGGCAAAAATA 123797  
QY 2522 CGACACGCGCCACACTGGGCGATGGGACACAGCAGCATGGAGCGAAAACAGTGCAAAATG 2581  
Db 123796 CGACACGCGCCACACTGGGCGATGGGACACAGCAGCATGGAGCGAAAACAGTGCAAAATG 123737  
QY 2582 CAAAAACCGCACAGCATAGTCTGTTTGCAGGCATACGGCACGATGCGGCGGATATCGGCT 2641  
Db 123736 CAAAAACCGCACAGCATAGTCTGTTTGCAGGCATACGGCACGATGCGGCGGATATCGGCT 123677  
QY 2642 ATCTCAAAGCGCTGTTCTCCTACGGACGCTACAAAACAGCATACGCGCAGCAGCGGTG 2701  
Db 123676 ATCTCAAAGCGCTGTTCTCCTACGGACGCTACAAAACAGCATACGCGCAGCAGCGGTG 123617  
QY 2702 CGSACGAACATCGGAAGCAGCGTCAACGGCAGCGTATGACGCTGGCGGCACTGGCGG 2761  
Db 123616 CGSACGAACATCGGAAGCAGCGTCAACGGCAGCGTATGACGCTGGCGGCACTGGCGG 123557  
QY 2762 GTGTCAACGTTCCGTTTTCGCAACCGGGAGATTTGACGGTTCGAAGCGGCTCTCGGCTACG 2821  
Db 123556 GTGTCAACGTTCCGTTTTCGCAACCGGGAGATTTGACGGTTCGAAGCGGCTCTCGGCTACG 123497  
QY 2822 ACCTGCTCAAAACAGTATGCTTCGCCGCAAAAAGCAGTGTCTTTGGCTGGAGCGCAACA 2881  
Db 123496 ACCTGCTCAAAACAGTATGCTTCGCCGCAAAAAGCAGTGTCTTTGGCTGGAGCGCAACA 123437  
QY 2882 GCCTCACTGAAGCAGCAGTGGTTCGGACTCGCGGGTCTGAAGCTGTGCAACCCCTTGAGCG 2941  
Db 123436 GCATCACTGAAGCAGCAGTGGTTCGGACTCGCGGGTCTGAAGCTGTGCAACCCCTTGAGCG 123377  
QY 2942 ATAAACCGTCTCTGTTTGAACCGCGGGCTGGAACCGCAACCTGAAACGGACGCGACTACA 3001  
Db 123376 ATAAACCGTCTCTGTTTGAACCGCGGGCTGGAACCGCAACCTGAAACGGACGCGACTACA 123317  
QY 3002 CGGTAAACGGCGCGCTTTTACCGCGCGAGCTGCAGCAACCGGCAAGAGCGGGGCAACGAATA 3061  
Db 123316 CGGTAAACGGCGCGCTTTTACCGCGCGAGCTGCAGCAACCGGCAAGAGCGGGGCAACGAATA 123257  
QY 3062 TGCCGCGACACCCCGCTGTTTTCGGGCTGGCGCGGATGTCGAATTCGGCAACAGCGCTGGA 3121  
Db 123256 TGCCGCGACACCCCGCTGTTTTCGGGCTGGCGCGGATGTCGAATTCGGCAACAGCGCTGGA 123197

QY 3122 ACGGCTTGGCAGCTTACAGCTACGCCGGTTCCAAAACAGTACGGCAACACACAGCGGACGAG 3181  
|||||  
Db 123196 ACGGCTTGGCAGCTTACAGCTACGCCGGTTCCAAAACAGTACGGCAACACACAGCGGACGAG 123137  
|||||  
QY 3182 TCGGCGTAGCTACCGGTTCTGA 3204  
|||||  
Db 123136 TCGGCGTAGCTACCGGTTCTGA 123114  
|||||  
RESULT 4  
AX236469 4218 bp DNA linear PAT 26-SEP-2001  
LOCUS  
DEFINITION Sequence 162 from Patent WO0164922.  
ACCESSION AX236469  
VERSION AX236469.1 GI:15796056  
KEYWORDS  
SOURCE synthetic construct.  
ORGANISM synthetic construct.  
REFERENCE 1 (bases 1 to 4218)  
AUTHORS Arico,M.B., Comanducci,M.C., Galeotti,C.C., Masignani,V.C.,  
Guiliani,M.M. and Pizzi,M.C.  
TITLE Heterologous expression of neisserial proteins  
JOURNAL Patent: WO 0164922-A 162 07-SEP-2001;  
Chiron Spa (IT)  
FEATURES  
Location/Qualifiers  
source 1..4218  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="961cl-983"  
BASE COUNT 1196 a 1149 c 1130 g 743 t  
ORIGIN  
Query Match 26.0%; Score 834; DB 6; Length 4218;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 1234; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 1963 CTGCTGGCTTCCCTCGACAGCGTCGAAAAACAGCGGGCAGTGAAGCGCACGCTGTGCC 2022  
|||||  
Db 2971 CTGCTGGCTTCCCTCGACAGCGTCGAAAAACAGCGGGCAGTGAAGCGCACGCTGTGCC 3030  
QY 2023 TATTATGCTCGGGGAATGCGGCACGAGTCTTCGGCAGCGGCACATTCGCGGCC 2082  
|||||  
Db 3031 TATTATGCTCGGGGAATGCGGCACGAGTCTTCGGCAGCGGCACATTCGCGGCC 3090  
QY 2083 GCGGCTCTGAACACGCCCTAGAACAGGCGGCAGCAATCTGAAAAACCTGATGGTCGAA 2142  
|||||  
Db 3091 GCGGCTCTGAACACGCCCTAGAACAGGCGGCAGCAATCTGAAAAACCTGATGGTCGAA 3150  
QY 2143 CTGGATGCTCCGGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCGCGCCGACCGC 2202  
|||||  
Db 3151 CTGGATGCTCCGGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCGCGCCGACCGC 3210  
QY 2203 ACAGATATGCGGGGATCCGCCCTACAGCGGCACTTCGCGGCGAGCGGCGGCTACAG 2262  
|||||  
Db 3211 ACAGATATGCGGGGATCCGCCCTACAGCGGCACTTCGCGGCGAGCGGCGGCTACAG 3270  
QY 2263 CATGGAATGCGCGGACGCTGACGATCTCAACAGTCTCGCGCTACCGTCTATGCC 2322  
|||||  
Db 3271 CATGGAATGCGCGGACGCTGACGATCTCAACAGTCTCGCGCTACCGTCTATGCC 3330  
QY 2323 GACAGTACCGCGCCCATGCCGATATGAGGACGCGCGCTGAAAGCCGCTATCGGACGG 2382  
|||||  
Db 3331 GACAGTACCGCGCCCATGCCGATATGAGGACGCGCGCTGAAAGCCGCTATCGGACGG 3390  
QY 2383 TTGGACCAACAGCTACGGGTCTGGCGTCTATCGGCAACCAACAGACGCGTGAACG 2442  
|||||  
Db 3391 TTGGACCAACAGCGGCGGTCTGAAGGCAAAATGCGCGGCGAGTACCCAAACCGTTCGCG 3450  
QY 2443 TGGGACAGGCGGTCTTGAAGGCAAAATGCGCGGCGAGTACCCAAACCGTTCGCGCATGCC 2502  
|||||  
Db 3451 TGGGACAGGCGGTCTTGAAGGCAAAATGCGCGGCGAGTACCCAAACCGTTCGCGCATGCC 3510

QY 2503 GCGAAACCGCGGAAATATACAGCAGCGCGCCACACTGGGCATGGGACACAGCATGG 2562  
|||||  
Db 3511 GCGAAACCGCGGAAATATACAGCAGCGCGCCACACTGGGCATGGGACACAGCATGG 3570  
|||||  
QY 2563 AGCGAAACAGTGCATAATGCAAAACCGCAGCAGATTAGTCTGTTTGGAGGCATACGGCAC 2622  
|||||  
Db 3571 AGCGAAACAGTGCATAATGCAAAACCGCAGCAGATTAGTCTGTTTGGAGGCATACGGCAC 3630  
|||||  
QY 2623 GATCGGGGCGATATCGCTATCTCAAAAGGCTCTTCTCTACGAGCGCTACAAAAACAGC 2682  
|||||  
Db 3631 GATCGGGGCGATATCGCTATCTCAAAAGGCTCTTCTCTACGAGCGCTACAAAAACAGC 3690  
QY 2683 ATCAGCCGACGACCGCTGCGGACGAACATGCGGAAGGACGCTCAACGCGACCGCTGATG 2742  
|||||  
Db 3691 ATCAGCCGACGACCGCTGCGGACGAACATGCGGAAGGACGCTCAACGCGACCGCTGATG 3750  
QY 2743 CAGCTGGGCGCACTGGGCGGTGTCAAGTTCCTTTTCCGCAACGGGAGATTGTACGGTC 2802  
|||||  
Db 3751 CAGCTGGGCGCACTGGGCGGTGTCAAGTTCCTTTTCCGCAACGGGAGATTGTACGGTC 3810  
QY 2803 GAAGCGGCTCTGGCTACGACCTGCTCAACAGGATGCTATCCGCCGAAAAAGCAGTGTCT 2862  
|||||  
Db 3811 GAAGCGGCTCTGGCTACGACCTGCTCAACAGGATGCTATCCGCCGAAAAAGCAGTGTCT 3870  
QY 2863 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGTGTCGAGCTCGCGGCTCTGAAG 2922  
|||||  
Db 3871 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGTGTCGAGCTCGCGGCTCTGAAG 3930  
QY 2923 CTGTCCCAACCTTTGAGCGGATTAAGCGCTCTTTTGAACGCGCGGCTGTGAACGCGGAC 2982  
|||||  
Db 3931 CTGTCCCAACCTTTGAGCGGATTAAGCGCTCTTTTGAACGCGCGGCTGTGAACGCGGAC 3990  
QY 2983 CTGAACGCGGCGACTACAGGTAACGGCGGCTTTTACCGCGCGGACTGCGCAACCGGC 3042  
Db 3991 CTGAACGCGGCGACTACAGGTAACGGCGGCTTTTACCGCGCGGACTGCGCAACCGGC 4050  
QY 3043 AAGACGGGCGACGCAATATGCGGCACACCGCGCTGTTGCCGCTGTGGGCGCGGATGTC 3102  
Db 4051 AAGACGGGCGACGCAATATGCGGCACACCGCGCTGTTGCCGCTGTGGGCGCGGATGTC 4110  
QY 3103 GAATTCGGCAACGGCTGGGAACGGCTTTGGCAGCTTACAGCTACGCCGCTTCCAAACAGTAC 3162  
Db 4111 GAATTCGGCAACGGCTGGGAACGGCTTTGGCAGCTTACAGCTACGCCGCTTCCAAACAGTAC 4170  
QY 3163 GGCACACACAGCGGCGAGTGGCGGCTAGGCTACCGGTTCTGA 3204  
Db 4171 GGCACACACAGCGGCGAGTGGCGGCTAGGCTACCGGTTCTGA 4212  
RESULT 5  
AX239755 4218 bp DNA linear PAT 26-SEP-2001  
LOCUS  
DEFINITION Sequence 57 from Patent WO0164920.  
ACCESSION AX239755  
VERSION AX239755.1 GI:15797388  
KEYWORDS  
SOURCE synthetic construct.  
ORGANISM synthetic construct.  
REFERENCE 1 (bases 1 to 4218)  
AUTHORS Arico,M.B., Comanducci,M., Galeotti,C., Masignani,V., Guiliani,M.M.  
and Pizzi,M.  
TITLE Hybrid expression of neisserial proteins  
JOURNAL Patent: WO 0164920-A 57 07-SEP-2001;  
Chiron Spa (IT)  
FEATURES  
Location/Qualifiers  
source 1..4218  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="961cl-983"  
BASE COUNT 1196 a 1149 c 1130 g 743 t  
ORIGIN



Query Match		26.0%;	Score 834;	DB 6;	Length 4218;
Best Local Similarity		99.4%;	Pred. No. 0;		
Matches 1234;		Conservative 0;	Mismatches 8;	Indels 0;	Gaps 0;
Qy	1963	CTGCTGGCTTCCTCGACAGCGTCCGAAACACAGCGGACGAGTGAAGGCGACACGCTGTGCC	2022		
Db	2971	CTGCTGGCTTCCTCGACAGCGTCCGAAACACAGCGGCGAGTGAAGGCGACACGCTGTGCC	3030		
Qy	2023	TATTATGTCCTGCGCGGAATCGGCACGGACTGCTTCGGGACGCGGCACATTCGCGGCC	2082		
Db	3031	TATTATGTCCTGCGCGGAATCGGCACGGACTGCTTCGGGACGCGGCACATTCGCGGCC	3090		
Qy	2083	GCCGGTCTGAACACCGCGTAGAACAGGCGCGCAATCTGTGAAACCTGATGTGCA	2142		
Db	3091	GCCGGTCTGAACACCGCGTAGAACAGGCGCGCAATCTGTGAAACCTGATGTGCA	3150		
Qy	2143	CTGGATGCTCCGATCATCCGCAACACCCGAGACGGTTGAACCTGCGGCCGCCACCGC	2202		
Db	3151	CTGGATGCTCCGATCATCCGCAACACCCGAGACGGTTGAACCTGCGGCCGCCACCGC	3210		
Qy	2203	ACAGATATGCCGGGCATCGCCCTACGGCGCAACTTTCGGCGGACGCGCGCTACAG	2262		
Db	3211	ACAGATATGCCGGGCATCGCCCTACGGCGCAACTTTCGGCGGACGCGCGCTACAG	3270		
Qy	2263	CATGCGAATGCCGCGCGATGCTACGCATCTTCAACAGTCTCGCGCTACCGCTCATGCC	2322		
Db	3271	CATGCGAATGCCGCGCGATGCTACGCATCTTCAACAGTCTCGCGCTACCGCTCATGCC	3330		
Qy	2323	GACAGTACCGCGCCCATCGCATATGAGGAGACCGCGCTGAAGCCGTATCGACGGG	2382		
Db	3331	GACAGTACCGCGCCCATCGCATATGAGGAGACCGCGCTGAAGCCGTATCGACGGG	3390		
Qy	2383	TTGGACCAACAGCTACGGGTTCGCGCTCATCGCGCAAAACCCACAGACGGTGGAAAG	2442		
Db	3391	TTGGACCAACAGCGCAGCTTCGCGCTCATCGCGCAAAACCCACAGACGGTGGAAAG	3450		
Qy	2443	TGGGAACAGGCGGTGTGAAGGCAAAATGCGCGGACGATACCCAAACCGTCGGCATTTGCC	2502		
Db	3451	TGGGAACAGGCGGTGTGAAGGCAAAATGCGCGGACGATACCCAAACCGTCGGCATTTGCC	3510		
Qy	2503	CGGAAACCGGCAAAATACAGACAGCAGCGCCACATCGGCATGGACGACAGCACATGG	2562		
Db	3511	CGGAAACCGGCAAAATACAGACAGCAGCGCCACATCGGCATGGACGACAGCACATGG	3570		
Qy	2563	AGCGAAACAGTGCAAATGCAAAACCCGACAGCATTAGTCTCTTTTCAGGACATACGGAC	2622		
Db	3571	AGCGAAACAGTGCAAATGCAAAACCCGACAGCATTAGTCTCTTTTCAGGACATACGGAC	3630		
Qy	2623	GATGCGGCGATATCGGCTATCTCAAAGCCCTGTTCTCTACGGACGCTACAAAACAGC	2682		
Db	3631	GATGCGGCGATATCGGCTATCTCAAAGCCCTGTTCTCTACGGACGCTACAAAACAGC	3690		
Qy	2683	ATCAGCCGACGACCGGTGCGGACCAACATCGGGAAGCAGCGTCAACGGCAGCTGATG	2742		
Db	3691	ATCAGCCGACGACCGGTGCGGACCAACATCGGGAAGCAGCGTCAACGGCAGCTGATG	3750		
Qy	2743	CAGCTGGCGCACTGGCGGTGTCAACCTTCGGTTGCGCAACCGGAGATTGACGGTC	2802		
Db	3751	CAGCTGGCGCACTGGCGGTGTCAACCTTCGGTTGCGCAACCGGAGATTGACGGTC	3810		
Qy	2803	GAAGCGGTCTGCGGTACGACCTGTCAACAGGATGATTCGCCGCAAAAGGCGAGTCT	2862		
Db	3811	GAAGCGGTCTGCGGTACGACCTGTCAACAGGATGATTCGCCGCAAAAGGCGAGTCT	3870		
Qy	2863	TTGGGCTGAGCGGCAACAGCCTCACTGAAGGCACACTGTTGGGACTCGCGGTCTGAAG	2922		
Db	3871	TTGGGCTGAGCGGCAACAGCCTCACTGAAGGCACCGTGTTCGGACTCGCGGTCTGAAG	3930		
Qy	2923	CTGTCCCAACCCCTTGAGCGATAAAGCCCTCTCTGTTTGAACGCGCGGCGTGAACCGCAC	2982		
Db	3931	CTGTCCCAACCCCTTGAGCGATAAAGCCCTCTCTGTTTGAACGCGCGGCGTGAACCGCAC	3990		
Qy	2983	CTGAACGACGCGACTACAGGTACGCGGCGGCTTTACGCGCGGCGACTTGCACCAACCGCG	3042		

Db	3991	CTGAACGGACGCGCACTACACGGTAACGGGGCGGCTTTACGGCGCGCACTGCAGCAACCGCG	4050	
QY	3043	AAGACGGGGACGCAATATGCCGACACACCGCGCTGTTGCCGGTCTGGCGCGGATGTC	3102	
Db	4051	AAGACGGGGACGCAATATGCCGACACACCGCGTCTGGTTCGCCGCGCTGGCGCGGATGTC	4110	
QY	3103	GAATTCGGCAACGGCTCGAACGGCTTTGGCAGCTTACAGCTACGCCCGTTCCAAACAGTAC	3162	
Db	4111	GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCCGTTCCAAACAGTAC	4170	
QY	3163	GGCAACCAACGGCGAGTCGGCGGTAGGCTACCGGTTCTGA	3204	
Db	4171	GGCAACCAACGGCGAGTCGGCGGTAGGCTACCGGTTCTGA	4212	
RESULT 6				
AE002545 10591 bp DNA linear BCT 25-MAY-2000				
LOCUS				
DEFINITION				
Neisseria meningitidis serogroup B strain MC58 section 187 of 206 of the complete genome.				
ACCESSION				
AE002545 AE002098				
VERSION				
AE002545.1 GI:7272729				
KEYWORDS				
SOURCE				
ORGANISM				
Neisseria meningitidis MC58.				
Neisseria meningitidis MC58				
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.				
1 (bases 1 to 10591)				
Tetelin,H., Saunders,N.J., Heidelberg,J., Jeffries,A.C., Nelson,K.E., Eisen,J.A., Ketchum,K.A., Hood,D.W., Peden,J.F., Dodson,R.J., Nelson,W.C., Gwinn,M.L., DeBoy,R., Peterson,J.D., Hickey,E.K., Haft,D.H., Salzberg,S.L., White,O., Fleischmann,R.D., Dougherty,B.A., Mason,T., Ciecko,A., Parksey,D.S., Blair,E., Cittone,H., Clark,E.B., Cotton,M.D., Utterback,T.R., Khouri,H., Qin,H., Vamathevan,J., Gill,J., Scarlato,V., Masignani,V., Pizsa,M., Grandi,G., Sun,L., Smith,H.O., Fraser,C.M., Moxon,E.R., Rappuoli,R. and Venter,J.C.				
Complete genome sequence of Neisseria meningitidis serogroup B strain MC58				
Science 287 (5459), 1809-1815 (2000)				
20175755				
10710307				
2 (bases 1 to 10591)				
Tetelin,H., Saunders,N.J., Heidelberg,J., Jeffries,A.C., Nelson,K.E., Eisen,J.A., Ketchum,K.A., Hood,D.W., Peden,J.F., Dodson,R.J., Nelson,W.C., Gwinn,M.L., DeBoy,R., Peterson,J.D., Hickey,E.K., Haft,D.H., Salzberg,S.L., White,O., Fleischmann,R.D., Dougherty,B.A., Mason,T., Ciecko,A., Parksey,D.S., Blair,E., Cittone,H., Clark,E.B., Cotton,M.D., Utterback,T.R., Khouri,H., Qin,H., Vamathevan,J., Gill,J., Scarlato,V., Masignani,V., Pizsa,M., Grandi,G., Sun,L., Smith,H.O., Fraser,C.M., Moxon,E.R., Rappuoli,R. and Venter,J.C.				
Direct Submission				
Submitted (17-MAR-2000) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA				
Location/Qualifiers				
1. 10591				
/organism="Neisseria meningitidis MC58"				
/strain="MC58"				
/db_xref="taxon:122586"				
/note="serogroup: B"				
548..3796				
/gene="NMB1969"				
548..3796				
/gene="NMB1969"				
/note="similar to SP:P31631 GB:U07788 GB:M62363 PID:507431 percent identity: 48.64; identified by sequence similarity; putative"				
/codon_start=1				
/transl_table=11				
/product="serotype-1-specific antigen, putative"				
/protein_id="AAF4298.1"				



```
/db_xref="GI:7227230"
/translation="MRTTPTFTKTKPTAMALAVATLSACLGGGGGTSAPDFNAG
GTGIGSNRATTAASVAGYIKNEMCKDRSMCLACRDDVAVTDRDAKINAPPNHLH
TGDFFNPNDAYKNLNLKPAIEAGYTGRCVEGIVDTGESVGSISPELYGRKEHYN
ENYKNYATYMRKEAPDEGGGKDIEASFDEAVEIETAKDTIRHYKEIGHIDLVSII
GGRSVDRPAGGTADPAUHMINTDETAKEMVAAIRNAWVKLGRGVRIVNNSPGT
TRAGTADLFOJANSEEOYRQALLDDYSGGDKTDEGIRLMOQSDYGNLSYHIRKNMLF
IFSTGNDAAQNTALLPFYEKDAQGIIITVAGVDRSGEKFREMYGPPEPLEYG
SNHGCTTAMWCLISAPYEASVFRTRNPIQIAGTSFAPITVGTAAALLQKYDPMWSDN
LRTLLTTAADDIGAGVDSKFCWGLLDACKAMNGPASFPFGDFTADTKGTSDIAYSNR
NDSITGGGLIKGGSQLGHNNTYTKLIEGGSLVLYGNKNSDMRVETKALINYNG
AASGSLNSGDIYLLADQGSQNETVHLKSLQDLGKTLIRUGLKLKLVDSGTALIG
KLYMSARGGAGYLNSTGRVPFFLSAAKIGQDYSFTNIEITDGLASLSDVETKAG
SEGDTLSYYVRNGRARTASAAHAPAGLKHAVEQGSNLENLWVELDASSTAPPE
TVETAAADRTDMPGIRPYCATFRAAAAOHANAADGVRIFNLSLAATVYADSTAAHADM
QGRRLKVASDGLDHNGTGLRVIAOTQDDGTWEOGVGCKMRGSTQTVGIKAKTGMT
TAATLGMGRSTWSENANAKTDSISLFIAGIRHDAGDIGLKLGSYGRYKNSISRST
GADEHAEGSVNGTLMQALGALGVNVPFAATGDLTVEGGLRDLLDKODAEKGSALGW
SGNSTEGTLVLGAGLKLSQLPSDKAVLAFATAGVERDLNGRDYTVTGFTGATAATGK
TGARNPHPTRLVAGLGADVEFGNGWNLARYSVAGSKOYGNHSGRVGVGYRF"
/length=4828
/genes="NMB1970"
4828..6609
/notes="similar to GB:AE000511 PID:2313385 percent
identity: 53.22; identified by sequence similarity;
putative"
/codon_start=1
/trans_table=11
/product="para-aminobenzoate synthetase component
1/4-amino-4-deoxychorismate lyase, putative"
/protein_id="AAF42299.1"
/db_xref="GI:7227231"
/translation="MPYFALFDADVSGRAKRYONHVESFFPEELDALDGLQSCWQ
KGLHSVLFADYGFGLPTGVESERGNLALHFWANCADDAESLWLRHSDGLPAGLST
PQSVSETDILRKQIHEAIRRGDTQIYNTIRLHLQAYGNPVSILRLRQPPVAV
LHLDPAEGCSAWTLCFSPFLKIGSDGTISTEPMKGAPILQDQDERRAEALQAD
PKNRAENVILLRNDLGLKIAOTGTCVPEPKVSFRGSMQMTSTIOAALPHITSF
ADILRAFPCCGSTITGAPKMSQIIESLEAEARGLYTGISYLNPCSGLGFEGFNV
VIRLSTLPLSDIYQGVYGVSGIVIDSDPAEAYRECCGKARFELNLRPDRGIFETL
RAENGRTLLDRHLCRLTSAOALNPLPDGCENQIKOYIADLPDGAFFVKALLASDG
ISLSRAVNLRLDKQVLIISPAVLPAQNYLRRFKTCRALFDQAMQTAETQGAFLSF
FNSDGLLEGGRSNFIKRRGQWLTPSLDLIDLINGIMRQAVLDEPOKYLOTNOVIETH
ITQKTLQEAEEIRLSNALRGVFAAALA"
/complement(6668..8185)
/genes="NMB1971"
complement(6668..8185)
/notes="conserved hypothetical protein; identified by
Glimmer2; putative"
/codon_start=1
/trans_table=11
/product="conserved hypothetical protein"
/protein_id="AAF42300.1"
/db_xref="GI:7227232"
/translation="MLYFRYFLVWCAAGVSAAYGADAPAILDDKALLQVORSVDK
WAESQKVNDAAPRVVDGDFLLAHPKMLEHSLRDALNGQADLIASLADLYAKLPDYD
AVLYGRARALLAKLAPGPAEAVARYRELHGENAADERIILDLAAAFDFDFRLKSAERH
PAEAKLDPAPLVENVGRFRKKTGELTGRFSGGISPAVRNANNAAPQYCRONGGR
QICSVRAARAGLNTIEAEKLTPLADNHYLLFRNIGNGTIYTFYSKSAYDDGFGRA
YLGWQYKNARQATGLIPFYQVQLSGSDGFDKTKRVNNRRLPPIYMLAHGVQVLSHTY
RNPQWQFVVALEHRYORYREQDRAEYNNRGDGFYVSSAKRLGESATVFGWQFVRF
VPKRETVGGAVNNAAYRRNGVYAGWAEWRQLGGLNSRVSAVARNYKGIASFTEA
QNRNENNVSLALSHDKLSYKGIVPALNYLREGRTESNVPVYKRNSEVFFVSADWR"
/complement(8881..10515)
/genes="NMB1972"
complement(8881..10515)
/notes="similar to SP:P29842 GB:223008 GB:223009 PID:438185
PID:438187 percent identity: 99.26; identified by sequence
similarity; putative"
/codon_start=1
/trans_table=11
```

```
/product="chaperonin, 60 kDa"
/protein_id="AAF42301.1"
/db_xref="GI:7227233"
/translation="MAAKDVQFGNEVRQKMWNVNLTANAVRVTLQPKGRNVVVDRAF
GGPHITKDGVTYVAKEIELDKFENMGAMQVKEVASKTNDVAGTGTITATVLAQSVIAE
GMYRTAGVNTPTDLKRGIDKAAVALDELKNTAKPCDTSKEIAQVGSISANSDEQVGA
ITAEMERKYGKEGIVTEDEGKSLNEDLVVEGQFDRGLYLSPIFIANDKQIATAALDNP
FVLLPDKITSNIRDLPLVEQVAKASRPILLIIAEDVEGALATLVNNIRGILKTVAV
KAPGFGDRKAMLODIAIITGGVVVISEEVGLSLEKATLDDLOGAKEFIEIGKNTTIIID
GFGDAQIAEARVAEIRQIETATSDYDKELQBERVAKLAGGVAVIKVGAATEVEMKEK
KORVEDAHATRAEIRVEGVVAGGVALLRAALENDLHGTGADQDQVQIVLRAVESP
LRQIVAGGEPSSVYVNVKLEGGKTYGNAGSEYGDYMIEMGVLDPAKVRTSALQHA
STAGLMLTTDCMIAIEIPDKPVDNMGGMGMGM"
/length=2623
/genes="NMB1970"
2623..3083
/notes="similar to GB:AE000511 PID:2313385 percent
identity: 53.22; identified by sequence similarity;
putative"
/codon_start=1
/trans_table=11
/product="para-aminobenzoate synthetase component
1/4-amino-4-deoxychorismate lyase, putative"
/protein_id="AAF42299.1"
/db_xref="GI:7227231"
/translation="MPYFALFDADVSGRAKRYONHVESFFPEELDALDGLQSCWQ
KGLHSVLFADYGFGLPTGVESERGNLALHFWANCADDAESLWLRHSDGLPAGLST
PQSVSETDILRKQIHEAIRRGDTQIYNTIRLHLQAYGNPVSILRLRQPPVAV
LHLDPAEGCSAWTLCFSPFLKIGSDGTISTEPMKGAPILQDQDERRAEALQAD
PKNRAENVILLRNDLGLKIAOTGTCVPEPKVSFRGSMQMTSTIOAALPHITSF
ADILRAFPCCGSTITGAPKMSQIIESLEAEARGLYTGISYLNPCSGLGFEGFNV
VIRLSTLPLSDIYQGVYGVSGIVIDSDPAEAYRECCGKARFELNLRPDRGIFETL
RAENGRTLLDRHLCRLTSAOALNPLPDGCENQIKOYIADLPDGAFFVKALLASDG
ISLSRAVNLRLDKQVLIISPAVLPAQNYLRRFKTCRALFDQAMQTAETQGAFLSF
FNSDGLLEGGRSNFIKRRGQWLTPSLDLIDLINGIMRQAVLDEPOKYLOTNOVIETH
ITQKTLQEAEEIRLSNALRGVFAAALA"
/complement(6668..8185)
/genes="NMB1971"
complement(6668..8185)
/notes="conserved hypothetical protein; identified by
Glimmer2; putative"
/codon_start=1
/trans_table=11
/product="conserved hypothetical protein"
/protein_id="AAF42300.1"
/db_xref="GI:7227232"
/translation="MLYFRYFLVWCAAGVSAAYGADAPAILDDKALLQVORSVDK
WAESQKVNDAAPRVVDGDFLLAHPKMLEHSLRDALNGQADLIASLADLYAKLPDYD
AVLYGRARALLAKLAPGPAEAVARYRELHGENAADERIILDLAAAFDFDFRLKSAERH
PAEAKLDPAPLVENVGRFRKKTGELTGRFSGGISPAVRNANNAAPQYCRONGGR
QICSVRAARAGLNTIEAEKLTPLADNHYLLFRNIGNGTIYTFYSKSAYDDGFGRA
YLGWQYKNARQATGLIPFYQVQLSGSDGFDKTKRVNNRRLPPIYMLAHGVQVLSHTY
RNPQWQFVVALEHRYORYREQDRAEYNNRGDGFYVSSAKRLGESATVFGWQFVRF
VPKRETVGGAVNNAAYRRNGVYAGWAEWRQLGGLNSRVSAVARNYKGIASFTEA
QNRNENNVSLALSHDKLSYKGIVPALNYLREGRTESNVPVYKRNSEVFFVSADWR"
/complement(8881..10515)
/genes="NMB1972"
complement(8881..10515)
/notes="similar to SP:P29842 GB:223008 GB:223009 PID:438185
PID:438187 percent identity: 99.26; identified by sequence
similarity; putative"
/codon_start=1
/trans_table=11
```

|||||  
Db 3335 CAGCTGGCGCAGCTGGCGGCTGCTCAACGTTCCGTTTCCGCCAACGGAGATTGACGGTC 3394  
QY 2803 GAAGGGGTCTGGCTTACCACCTGCTCAACAGGATGCAATCCGCCAAAGAGCGAGTGTCT 2862  
Db 3395 GAAGGGGTCTGGCTTACCACCTGCTCAACAGGATGCAATCCGCCAAAGAGCGAGTGTCT 3454  
QY 2863 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTGGACTCGCGGCTCTGAAG 2922  
Db 3455 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTGGACTCGCGGCTCTGAAG 3514  
QY 2923 CTGTCCCAACCTTTGAGCGATAAAGCCGCTCTGTTTGAACGGCGGCTGGAAACGCGAC 2982  
Db 3515 CTGTCCCAACCTTTGAGCGATAAAGCCGCTCTGTTTGAACGGCGGCTGGAAACGCGAC 3574  
QY 2983 CTGAACGGAGCGGACTTACACGCTTAACGGCGGCTTTTACCGCGGCGACTGCACGACCGGC 3042  
Db 3575 CTGAACGGAGCGGACTTACACGCTTAACGGCGGCTTTTACCGCGGCGACTGCACGACCGGC 3634  
QY 3043 AAGACGGGGCACGCAATATGCCGACACCCCGCTGTTTGGCGGCTGGCGCGGATGTC 3102  
Db 3635 AAGACGGGGCACGCAATATGCCGACACCCCGCTGTTTGGCGGCTGGCGCGGATGTC 3694  
QY 3103 GAATTCGGAACGGCTGGAAACGGCTTTGGCACGTTTACAGCTACGCCGGTTTCCAAACAGTAC 3162  
Db 3695 GAATTCGGAACGGCTGGAAACGGCTTTGGCACGTTTACAGCTACGCCGGTTTCCAAACAGTAC 3754  
QY 3163 GGCACACAGCGGAGCTGCGCGTAGCGTACCGGTTCTGA 3204  
Db 3755 GGCACACAGCGGAGCTGCGCGTAGCGTACCGGTTCTGA 3796

RESULT 7  
AX044034  
LOCUS AX044034 349980 bp DNA linear PAT 24-NOV-2000  
DEFINITION Sequence 113 from Patent WO0066791.  
ACCESSION AX044034  
VERSION AX044034.1 GI:11342918

KEYWORDS  
SOURCE  
ORGANISM  
Neisseria meningitidis.  
Neisseria meningitidis.  
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;  
Neisseria.  
REFERENCE  
AUTHORS 1 (bases 1 to 349980)  
Piazza, M., Hickey, E., Peterson, J., Tettelin, H., Venter, J.C.,  
Masignani, V., Galeotti, C., Mora, M., Ratti, G., Scarselli, M.,  
Scarlato, V., Rappuoli, R., Frazer, C.M. and Grandi, G.  
TITLE  
JOURNAL Neisseria genomic sequences and methods of their use  
Patent: WO 0066791-A 113 09-NOV-2000;  
CHIRON CORPORATION (US); THE INSTITUTE FOR GENOMIC RESEARCH (US)

FEATURES  
Location/Qualifiers  
Source  
1..349980  
/organism="Neisseria meningitidis"  
/db\_xref="taxon:487"  
/note="sequence too long, cut in 8 pieces.-seq 1: 1 to  
349980 349980 bases-seq 108: 300001 to 649980 349980  
bases-seq 109: 600001 to 949980 349980 bases-seq 110:  
900001 to 1249980 349980 bases-seq 111: 1200001 to 1549980  
349980 bases-seq 112: 1500001 to 1849980 349980 bases-seq  
113: 1800001 to 2149980 349980 bases-seq 114: 2100001 to  
2272325 172325 bases"  
BASE COUNT 86473 a 95646 c 85908 g 81953 t  
ORIGIN

Query Match 26.0%; Score 834; DB 6; Length 349980;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 1234; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 1963 CTGCTGGCTTCCCTCCACAGCGTCGAAAAACAGCGGGCAGTGAAGCGCACAGCGTGTCC 2022  
Db 263419 CTGCTGGCTTCCCTCCACAGCGTCGAAAAACAGCGGGCAGTGAAGCGCACAGCGTGTCC 263478  
QY 2023 TATTATGTCGCGGCAATGCGGACGCGACTGTCTCGGCAGCGCACATTCCTCCGCGCC 2082

|||||  
Db 263479 TATTATGTCGCGGCAATGCGGACGAGCTGTTGGCAGCGCACATTCCTCCGCGCC 263538  
QY 2083 GCCGCTCTGAACACGCGCTAGAACAGGCGGCACCAATCTGGAACAACTGATGTCGAA 2142  
Db 263539 GCCGCTCTGAACACGCGCTAGAACAGGCGGCAGCAATCTGGAACAACTGATGTCGAA 263598  
QY 2143 CTGGATGCTCCGAAATCATCCGCAACACCCGAGAGCGGTTGAAACTCGCGCCGCCACCGC 2202  
Db 263599 CTGGATGCTCCGAAATCATCCGCAACACCCGAGAGCGGTTGAAACTCGCGCACCGACCGC 263658  
QY 2203 ACAGATATGCCGGGCGATCCGCCCTTACGGCGCAACTTTCGGCGCAGCGCACCGGTACAG 2262  
Db 263659 ACAGATATGCCGGGCGATCCGCCCTTACGGCGCAACTTTCGGCGCAGCGCACCGGTACAG 263718  
QY 2263 CATGCGAATGCCCGCGAGCGGTGTACGCACTTTCACAGTCTCGCGCTACCGCTCTATGCC 2322  
Db 263719 CATGCGAATGCCCGCGAGCGGTGTACGCACTTTCACAGTCTCGCGCTACCGCTCTATGCC 263778  
QY 2323 GACAGTACCGCGCGCCATGCCGATATGAGGAGCGCGGCTGAAAGCCGTATCGGACGGG 2382  
Db 263779 GACAGTACCGCGCGCCATGCCGATATGAGGAGCGCGGCTGAAAGCCGTATCGGACGGG 263838  
QY 2383 TTGGACCAACAGCTTACGGGTCTGCGGTTCATCGCGCAAAACCAACAGAGCGGTGGAAGC 2442  
Db 263839 TTGGACCAACAGCGGCTTGGCGGTCTATCGCGCAAAACCAACAGAGCGGTGGAAGC 263898  
QY 2443 TGGGAACAGGCGGCTTGAAGGCAAAATGCGGCGAGTACCCAAACCGCTCGGCATTGGC 2502  
Db 263899 TGGGAACAGGCGGCTTGAAGGCAAAATGCGGCGAGTACCCAAACCGCTCGGCATTGGC 263958  
QY 2503 GCGAAACCGCGGCAAAATACGACAGCGCGCACACTGGGATGGGACACAGCACATGG 2562  
Db 263959 GCGAAACCGCGGCAAAATACGACAGCGCGCACACTGGGATGGGACAGCACATGG 264018  
QY 2563 AGCGAAACAGTGCATAATGCAAAACCGACAGCATTAGTCTTTTTCAGAGCATACGGCAC 2622  
Db 264019 AGCGAAACAGTGCATAATGCAAAACCGACAGCATTAGTCTTTTTCAGAGCATACGGCAC 264078  
QY 2623 GATGCGGCGATATCGGCTATCTCAAGGCGCTTCTCTACGAGCGCTTACAAACACG 2682  
Db 264079 GATGCGGCGATATCGGCTATCTCAAGGCGCTTCTCTACGAGCGCTTACAAACACG 264138  
QY 2683 ATCAGCGCGACGCGGTGCGGACGAAACATGCGGAGGAGCGGTCAACGCGCAGCTGATG 2742  
Db 264139 ATCAGCGCGACGCGGTGCGGACGAAACATGCGGAGGAGCGGTCAACGCGCAGCTGATG 264198  
QY 2743 CAGCTGGCGCACCTGGCGGTGTCAACGTTTCGTTTCCGCAACGGGAGATTGACGGTC 2802  
Db 264199 CAGCTGGCGCACCTGGCGGTGTCAACGTTTCGTTTCCGCAACGGGAGATTGACGGTC 264258  
QY 2803 GAAGGCGGTCTCGGCTACGACCTGCTCAACAGGATGCAATCCGCAAAAGCGAGTGTCT 2862  
Db 264259 GAAGGCGGTCTCGGCTACGACCTGCTCAACAGGATGCAATCCGCAAAAGCGAGTGTCT 264318  
QY 2863 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTGGACTCGCGGCTCTGAAG 2922  
Db 264319 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTGGACTCGCGGCTCTGAAG 264378  
QY 2923 CTGTCCCAACCTTTGAGCGGATTAAGCCGCTCTGTTTGAACGGCGGCGGTGGAACCGCAC 2982  
Db 264379 CTGTCCCAACCTTTGAGCGGATTAAGCCGCTCTGTTTGAACGGCGGCGGTGGAACCGCAC 264438  
QY 2983 CTGAACGGAGCGGACTTACACGCTTAACGGCGGCTTTTACCGCGCGGACTGCACCAACCGC 3042  
Db 264439 CTGAACGGAGCGGACTTACACGCTTAACGGCGGCTTTTACCGCGCGGACTGCACCAACCGC 264498  
QY 3043 AAGAGGGGGCACGCAATATGCCGACACCCCGCTGTTTGGCGGCTTGGCGCGGATGTC 3102  
Db 264499 AAGAGGGGGCACGCAATATGCCGACACCCCGCTGTTTGGCGGCTTGGCGCGGATGTC 264558  
QY 3103 GAATTCGGCAACGGCTGGAAACGGCTTGGCACGTTTACAGCTACCGCGGTTTCCAAACAGTAC 3162  
|||||

Db 264559 GAATTCGGCAACGGCTGGAAACGGCTTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 264618  
QY 3163 GCGAACACAGCGGACGAGCTGGCGTAGCTACCGGTTCTGA 3204  
|||||  
Db 264619 GCGAACACAGCGGACGAGCTGGCGTAGCTACCGGTTCTGA 264660  
|||||  
RESULT 8  
AX236409  
LOCUS AX236409 3939 bp DNA linear PAT 26-SEP-2001  
DEFINITION Sequence 102 from Patent WO0164922.  
ACCESSION AX236409  
VERSION AX236409.1 GI:15796027  
KEYWORDS  
SOURCE synthetic construct.  
ORGANISM synthetic construct.  
REFERENCE 1 (bases 1 to 3939)  
AUTHORS Arico,M.B., Comanducci,M.C., Galeotti,C.C., Masignani,V.C.,  
Guiliani,M.M. and Pizze,M.C.  
TITLE Heterologous expression of neisserial proteins  
JOURNAL Patent: WO 0164922-A 102 07-SEP-2001;  
Chiron Spa (IT)  
FEATURES  
source Location/Qualifiers  
1..3939  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="deltaG983-741"  
BASE COUNT 1046 a 1100 c 1107 g 686 t  
ORIGIN  
Query Match 25.9%; Score 831; DB 6; Length 3939;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 1231; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 1963 CTGCTGGCTTCCCTCGACAGCGTTCGAAAAACAGCGGGCAGTGAAGCGACACGCTGTCC 2022  
|||||  
Db 1906 CTGCTGGCTTCCCTCGACAGCGTTCGAAAAACAGCGGGCAGTGAAGCGACACGCTGTCC 1965  
|||||  
QY 2023 TATTATGTCGTCGCGCAATCGGCACGACTGCTTCGGCAGCGCACATTCGCGGCC 2082  
|||||  
Db 1966 TATTATGTCGTCGCGCAATCGGCACGACTGCTTCGGCAGCGCACATTCGCGGCC 2025  
|||||  
QY 2083 GCGGCTCTGAAACACGCCGTAGAACAGCGCGCAGCAATCTGAAAAACCTGTATGTCGAA 2142  
|||||  
Db 2026 GCGGCTCTGAAACACGCCGTAGAACAGCGCGCAGCAATCTGAAAAACCTGTATGTCGAA 2085  
|||||  
QY 2143 CTGGATGCTCCGAATCATCCGAAACACCCGAGACGGTTGAAACTGCGGCCGCCGCCGCCG 2202  
|||||  
Db 2086 CTGGATGCTCCGAATCATCCGAAACACCCGAGACGGTTGAAACTGCGGCCGCCGCCGCCG 2145  
|||||  
QY 2203 ACAGATATCGCGGGATCGCCCTACGCGCAACTTTCGGCGCAGCGCACGCGGTACAG 2262  
|||||  
Db 2146 ACAGATATCGCGGGATCGCCCTACGCGCAACTTTCGGCGCAGCGCACGCGGTACAG 2205  
|||||  
QY 2263 CATGGAATCGCGCGACGCGTGTACGATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2322  
|||||  
Db 2206 CATGGAATCGCGCGACGCGTGTACGATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2265  
|||||  
QY 2323 GACGATACCGCGCCGATCGCCGATATGACGAGACGCCGCTGAAAGCCGATATCGGACGGG 2382  
|||||  
Db 2266 GACGATACCGCGCCGATCGCCGATATGACGAGACGCCGCTGAAAGCCGATATCGGACGGG 2325  
|||||  
QY 2383 TTGGACCAACGCTACGGGCTCGCGCTCATCGGCCAACCCACAGACGCGTGAACG 2442  
|||||  
Db 2326 TTGGACCAACGCGACGGGCTCGCGCTCATCGGCCAACCCACAGACGCGTGAACG 2385  
|||||  
QY 2443 TGGGAACAGCGCGGTCTGAAGCAAAATGCGCGGAGTACCAACCGCTCGGCATTGCC 2502  
|||||  
Db 2386 TGGGAACAGCGCGGTCTGAAGCAAAATGCGCGGAGTACCAACCGCTCGGCATTGCC 2445  
|||||  
QY 2503 GCGAAACCGCGGAAATACGACAGCGCGCCACACTGGGCTGGGACACAGCATGG 2562  
|||||

Db 2446 GCGAAACCGCGGAAATACGACAGCAGCGCGCCACACTGGGCATGGGACGACGACATGG 2505  
QY 2563 ACGGAAACAGTGCATGCAATGCAAAACCGGACAGCATTTAGTCTGTTTTCAGGCATACGGCAC 2622  
|||||  
Db 2506 ACGGAAACAGTGCATGCAATGCAAAACCGGACAGCATTTAGTCTGTTTTCAGGCATACGGCAC 2565  
|||||  
QY 2623 GATGCGGCGGATATCGGCTATCTCAAAGGCTTCTCTCTACGAGCGCTACAAAAACAGC 2682  
|||||  
Db 2566 GATGCGGCGGATATCGGCTATCTCAAAGGCTTCTCTCTACGAGCGCTACAAAAACAGC 2625  
|||||  
QY 2683 ATCAGCGGACGACCGGTCGGGACGAACATGCGGAAGGAGCGGTCAACGCGCAGCTGATG 2742  
|||||  
Db 2626 ATCAGCGGACGACCGGTCGGGACGAACATGCGGAAGGAGCGGTCAACGCGCAGCTGATG 2685  
|||||  
QY 2743 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTTCGCGCAACGCGGAGATTTGACGGTC 2802  
|||||  
Db 2686 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTTCGCGCAACGCGGAGATTTGACGGTC 2745  
|||||  
QY 2803 GAAGCGGCTCTCGCTACGACCTGTCTCAAACAGGATGCAATTCGCCGAAAAAGGACGTGCT 2862  
|||||  
Db 2746 GAAGCGGCTCTCGCTACGACCTGTCTCAAACAGGATGCAATTCGCCGAAAAAGGACGTGCT 2805  
|||||  
QY 2863 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGCTCTGAAG 2922  
|||||  
Db 2806 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGCTCTGAAG 2865  
|||||  
QY 2923 CTGTCCCAACCTTTGAGCGATTAAGCGCTCTCTGTTTTCGAACGCGCGGCTGGAAACCGGAC 2982  
|||||  
Db 2866 CTGTCCCAACCTTTGAGCGATTAAGCGCTCTCTGTTTTCGAACGCGCGGCTGGAAACCGGAC 2925  
|||||  
QY 2983 CTGAACGCGACGCGACTACACGGTAACGGCGGCTTTTACCGCGCGACTGTCAGCAACCGGC 3042  
|||||  
Db 2926 CTGAACGCGACGCGACTACACGGTAACGGCGGCTTTTACCGCGCGACTGTCAGCAACCGGC 2985  
|||||  
QY 3043 AGACGGGCGACGCAATATGCGCACACCGCCTGTTTCGCGGTCTGGCGCGGATGTC 3102  
|||||  
Db 2986 AGACGGGCGACGCAATATGCGCACACCGCCTGTTTCGCGGCTGGCGCGGATGTC 3045  
|||||  
QY 3103 GAATTCGGCAACGGCTGGAAACGGCTTGGCAGCTTACAGTACCGCGGTTCCAAACAGTAC 3162  
|||||  
Db 3046 GAATTCGGCAACGGCTGGAAACGGCTTGGCAGCTTACAGTACCGCGGTTCCAAACAGTAC 3105  
|||||  
QY 3163 GCGAACACAGCGGACGAGTTCGGCGGTAGCTACCGGTTTC 3201  
|||||  
Db 3106 GCGAACACAGCGGACGAGTTCGGCGGTAGCTACCGGTTTC 3144  
|||||

RESULT 9  
AX236419  
LOCUS AX236419 3939 bp DNA linear PAT 26-SEP-2001  
DEFINITION Sequence 112 from Patent WO0164922.  
ACCESSION AX236419  
VERSION AX236419.1 GI:15796032  
KEYWORDS  
SOURCE synthetic construct.  
ORGANISM synthetic construct.  
REFERENCE 1 (bases 1 to 3939)  
AUTHORS Arico,M.B., Comanducci,M.C., Galeotti,C.C., Masignani,V.C.,  
Guiliani,M.M. and Pizze,M.C.  
TITLE Heterologous expression of neisserial proteins  
JOURNAL Patent: WO 0164922-A 112 07-SEP-2001;  
Chiron Spa (IT)  
FEATURES  
source Location/Qualifiers  
1..3939  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="deltaG741-983"  
BASE COUNT 1046 a 1103 c 1106 g 684 t  
ORIGIN

Query Match 25.9%; Score 831; DB 6; Length 3939;  
Best Local Similarity 99.4%; Pred. No. 0;

Matches 1231: Conservative 0; Mismatches 8; Indels 0; Gaps 0;									
Qy	1963	CTGCTGGCTTCCCTCGACAGCGTTCGAAAAACAGCGGCAGTGAAGGCGACACGCTGTGCC	2022						
Db	2674	CTGCTGGCTTCCCTCGACAGCGTTCGAAAAACAGCGGCAGTGAAGGCGACACGCTGTGCC	2733						
Qy	2023	TATTATGTCCTGCGGGCAATGCGGCAGCGACTGCTTCGGCAGCGGCACATTCGCGCGCC	2082						
Db	2734	TATTATGTCCTGCGGGCAATGCGGCAGCGACTGCTTCGGCAGCGGCACATTCGCGCGCC	2793						
Qy	2083	GCCGCTCTGAACACGCGGTAGAACAGGCGGCACCAATCTGGAACACCTGATGTCGAA	2142						
Db	2794	GCCGCTCTGAACACGCGGTAGAACAGGCGGCACCAATCTGGAACACCTGATGTCGAA	2853						
Qy	2143	CTGGATGCTCCGGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGCGCGCGACCGC	2202						
Db	2854	CTGGATGCTCCGGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGCGCGCGACCGC	2913						
Qy	2203	ACAGATATGCCGGGATCCGCCCTTACGGCGCAACTTTCGGCGCAGCGGCAGCTACAG	2262						
Db	2914	ACAGATATGCCGGGATCCGCCCTTACGGCGCAACTTTCGGCGCAGCGGCAGCTACAG	2973						
Qy	2263	CATGCGAATGCCGCGAGCGGTGACGATCTTCAACAGTCTCGCGCTACGCTCTATGCC	2322						
Db	2974	CATGCGAATGCCGCGAGCGGTGACGATCTTCAACAGTCTCGCGCTACGCTCTATGCC	3033						
Qy	2323	GACAGTACCGCGCCATATGCCGATATGACGAGCAGCGGCTGAAAGCGGTATCGGACGG	2382						
Db	3034	GACAGTACCGCGCCATATGCCGATATGACGAGCAGCGGCTGAAAGCGGTATCGGACGG	3093						
Qy	2383	TTGGACCAACACGCTTACGGGCTTGGCGCTCATCGCGCAAAACCCCAACAGGACGGTGAACG	2442						
Db	3094	TTGGACCAACACGCGGCTTGGCGCTCATCGCGCAAAACCCCAACAGGACGGTGAACG	3153						
Qy	2443	TGGGAACAGGCGGTGTTGAAGGCAAAATGCGCGCAGTACCCAAACCGTGGCAATTGCC	2502						
Db	3154	TGGGAACAGGCGGTGTTGAAGGCAAAATGCGCGCAGTACCCAAACCGTGGCAATTGCC	3213						
Qy	2503	GCGAAACCGCGGAAATACGACAGCAGCGCGCACACTGGGCATGGGACAGCACATGG	2562						
Db	3214	GCGAAACCGCGGAAATACGACAGCAGCGCGCACACTGGGCATGGGACAGCACATGG	3273						
Qy	2563	AGCGAAACAGTGAATGCAAAACCCGACAGCAATAGTCTGTTTGCAGGATACGGCAC	2622						
Db	3274	AGCGAAACAGTGAATGCAAAACCCGACAGCAATAGTCTGTTTGCAGGATACGGCAC	3333						
Qy	2623	GATCGGGCGATATCGGCTATCTCAAGGCTGTCTCTACGGACGCTACAAAACAGC	2682						
Db	3334	GATCGGGCGATATCGGCTATCTCTAAGGCTGTCTCTACGGACGCTACAAAACAGC	3393						
Qy	2683	ATCAGCCGACGACCGGTGCGGACGAACATCGGAAGGACGCGTCAACGGCAGCTGATG	2742						
Db	3394	ATCAGCCGACGACCGGTGCGGACGAACATCGGAAGGACGCGTCAACGGCAGCTGATG	3453						
Qy	2743	CAGCTGGGCGACTGGGCGGTGTCAACGTTCCGTTTCCGCAACGGGAGATTGACGGTC	2802						
Db	3454	CAGCTGGGCGACTGGGCGGTGTCAACGTTCCGTTTCCGCAACGGGAGATTGACGGTC	3513						
Qy	2803	GAAGCGGTCTGCGCTACGACTGCTCAACAGGATGATTCGCGGAAAGGCGAGTCT	2862						
Db	3514	GAAGCGGTCTGCGCTACGACTGCTCAACAGGATGATTCGCGGAAAGGCGAGTCT	3573						
Qy	2863	TTGGCTGAGCGGCAACAGCCTCACTGAAGGACACACTGGTTCGGAATCGCGGGTCTGAAG	2922						
Db	3574	TTGGCTGAGCGGCAACAGCCTCACTGAAGGACACACTGGTTCGGAATCGCGGGTCTGAAG	3633						
Qy	2923	CTGTGCGCAACCTTTAGCGGATAAAGCCGTCCTGTTTTCGCAACGGGGCGGTGGAACGCGAC	2982						
Db	3634	CTGTGCGCAACCTTTAGCGGATAAAGCCGTCCTGTTTTCGCAACGGGGCGGTGGAACGCGAC	3693						
Qy	2983	CTGAACGACCGGACTACACGCTAACGGCGGCTTTACGGCGGCGACTGCAGCAACCGGC	3042						
Db	3694	CTGAACGACCGGACTACACGCTAACGGCGGCTTTACGGCGGCGACTGCAGCAACCGGC	3753						

QY	3043	AAGACGGGGGACGCAATATGCCGCACACCGCGCTGTTGCCGGTCTGGGGCGCGATGTC	3102																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																</
----	------	--	------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	----

Qy	2443	TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGGGCAGTACCCAAACCGTCGGCATTTGCC	2502
Db	2386	TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGGGCAGTACCCAAACCGTCGGCATTTGCC	2445
Qy	2503	GCGAAACCGCGGAAATACGACAGCAGCGGCACACTGGGCGATGGGACACAGCACATGG	2562
Db	2446	GCGAAACCGCGGAAATACGACAGCAGCGGCACACTGGGCGATGGGACAGCACATGG	2505
Qy	2563	AGCGAAACAGTGCAAATGCAAAACCGACAGCATTTAGTCTGTTGCGAGGCATACGGCAC	2622
Db	2506	AGCGAAACAGTGCAAATGCAAAACCGACAGCATTTAGTCTGTTGCGAGGCATACGGCAC	2565
Qy	2623	GATCGGGCGATATCGGCTATCTCAAGGCGCTGTTCTCTACGACGCTACAAAAACAGC	2682
Db	2566	GATCGGGCGATATCGGCTATCTCAAGGCGCTGTTCTCTACGACGCTACAAAAACAGC	2625
Qy	2683	ATCAGCGCAGCAGCGGTGCGGAGCAACATGCGGAAGGCGTCAACAGCGCATTTGACGGTC	2742
Db	2626	ATCAGCGCAGCAGCGGTGCGGAGCAACATGCGGAAGGCGTCAACAGCGCATTTGACGGTC	2685
Qy	2743	CAGTGGGCGCACTGGGCGGTGTCAAGTTCGGTTCGCCCAACGGGAGATTTGACGGTC	2802
Db	2686	CAGTGGGCGCACTGGGCGGTGTCAAGTTCGGTTCGCCCAACGGGAGATTTGACGGTC	2745
Qy	2803	GAAGGCGGTCTGGCTACGACCTGCTCAAAACAGGATTCGCGGCAAAAGGCGAGTGCT	2862
Db	2746	GAAGGCGGTCTGGCTACGACCTGCTCAAAACAGGATTCGCGGCAAAAGGCGAGTGCT	2805
Qy	2863	TTGGCTGGAGCGGCAACAGCCCTACATGGAAGGCACACTGTGCGGATTCGGGGTCTGAAG	2922
Db	2806	TTGGCTGGAGCGGCAACAGCCCTACATGGAAGGCACACTGTGCGGATTCGGGGTCTGAAG	2865
Qy	2923	CTGTGCGCAACCTTGAGCGATAAGCGTCTCTTTGCAACGGCGGGCGTGGACCGGAC	2982
Db	2866	CTGTGCGCAACCTTGAGCGATAAGCGTCTCTTTGCAACGGCGGGCGTGGACCGGAC	2925
Qy	2983	CTGAACGGACCGGACTACCGGTAACGGCGGGCTTTACCGCGCGGACTGCGAGCAACCGGC	3042
Db	2926	CTGAACGGACCGGACTACCGGTAACGGCGGGCTTTACCGCGCGGACTGCGAGCAACCGGC	2985
Qy	3043	AAGACGGGCGCAGCAATATGCGGCACACCCCGCTGTTGCGGGTCTGGGCGGGATGTC	3102
Db	2986	AAGACGGGCGCAGCAATATGCGGCACACCCCGCTGTTGCGGGTCTGGGCGGGATGTC	3045
Qy	3103	GAATTGCGCAACCGCTGGAACGCTTGGCACGTTACAGCTACCGCGTTCCTCAACAGTAC	3162
Db	3046	GAATTGCGCAACCGCTGGAACGCTTGGCACGTTACAGCTACCGCGTTCCTCAACAGTAC	3105
Qy	3163	GGCAACACAGCGGACGAGTGGCGGTAGGCTACCGGTTCC 3201	
Db	3106	GGCAACACAGCGGACGAGTGGCGGTAGGCTACCGGTTCC 3144	

RESULT 11  
AX239729  
LOCUS AX239729  
DEFINITION Sequence 31 from Patent WO0164920.  
ACCESSION AX239729  
VERSION AX239729.1 GI:15797375  
KEYWORDS .  
SOURCE synthetic construct.  
ORGANISM synthetic construct.  
REFERENCE 1 (bases 1 to 3939).  
AUTHORS Arico,M.B., Comanducci,M., Galeotti,C., Masignani,V., Giuliani,M.M.  
and Pizza,M.  
TITLE Hybrid expression of neisserial proteins  
JOURNAL Patent: WO 0164920-A 31 07-SEP-2001;  
Chiron Spa (IT)  
FEATURES Location/Qualifiers  
source 1..3939  
/organism="synthetic construct"

		/db_xref="taxon:32630" /note="deltaG741-983"			
BASE COUNT		1046 a	1103 c	1106 g	684 t
ORIGIN					
Query Match		25.9%;		Score 831;	DB 6;
Best Local Similarity		99.4%;		Pred. No. 0;	
Matches 1231;		Conservative 0;		Mismatches 8;	Indels 0;
				Gaps 0;	
Qy	1963	CTGCTGCTTCCCTCGACAGCGTGCAGAAAAACACAGCGGGCAGTGAAGCGACACGCTGTCC	2022		
Db	2674	CTGCTGCTTCCCTCGACAGCGTGCAGAAAAACACAGCGGGCAGTGAAGCGACACGCTGTCC	2733		
Qy	2023	TATTATCTCGTTCGCGCAATGCGGCACGGACTGCTTTCGCGAGCGGCACATTCGCGCCCC	2082		
Db	2734	TATTATCTCGTTCGCGCAATGCGGCACGGACTGCTTTCGCGAGCGGCACATTCGCGCCCC	2793		
Qy	2083	GCCGGTCTGAACACACGCCGTAGAACAGGGCGGAGCAATCTGGAACACCTGATGTCGAA	2142		
Db	2794	GCCGGTCTGAACACACGCCGTAGAACAGGGCGGAGCAATCTGGAACACCTGATGTCGAA	2853		
Qy	2143	CTGATGCTTCGATCATCGCAACACCGGAGACGGTTGAAACTGGGCGCGGACCGC	2202		
Db	2854	CTGATGCTTCGATCATCGCAACACCGGAGACGGTTGAAACTGGGCGCGGACCGC	2913		
Qy	2203	ACAGATATCGGGCATCCGCCCTACGGCGCACTTTCGCGCAGCGGAGCGGTACAG	2262		
Db	2914	ACAGATATCGGGCATCCGCCCTACGGCGCACTTTCGCGCAGCGGAGCGGTACAG	2973		
Qy	2263	CATGCGAATGCCCGCAGGTGTACGCATCTTCAACAGTCTCGCGCGTACCGTCTATGCC	2322		
Db	2974	CATGCGAATGCCCGCAGGTGTACGCATCTTCAACAGTCTCGCGCGTACCGTCTATGCC	3033		
Qy	2323	GACATACCGCCCGCATGCGATATGCGGAGCGCGCGGTGAAGCCGTATGGAGCGG	2382		
Db	3034	GACATACCGCCCGCATGCGGATATGCGGAGCGCGCGGTGAAGCCGTATGGAGCGG	3093		
Qy	2383	TTGGACCACAAACGCTACGGGTCTGCGCGTCATCGCGCAAAACCCAAACAGGACGGTGAACG	2442		
Db	3094	TTGGACCACAAACGCTACGGGTCTGCGCGTCATCGCGCAAAACCCAAACAGGACGGTGAACG	3153		
Qy	2443	TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGGGCAGTACCCAAACCGTCGGCATTTGCC	2502		
Db	3154	TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGGGCAGTACCCAAACCGTCGGCATTTGCC	3213		
Qy	2503	GCGAAACCGCGGAAATACGACAGCAGCGCGGCACACTGGGCGATGGGACACAGCATGG	2562		
Db	3214	GCGAAACCGCGGAAATACGACAGCAGCGCGGCACACTGGGCGATGGGACACAGCATGG	3273		
Qy	2563	AGCGAAACAGTGCAAATGCAAAACCGACAGCATTTAGTCTGTTTTCGAGGCATACGGCAC	2622		
Db	3274	AGCGAAACAGTGCAAATGCAAAACCGACAGCATTTAGTCTGTTTTCGAGGCATACGGCAC	3333		
Qy	2623	GATCGGGCGATATCGGCTATCTCAAGGCGCTGTTCTCTACGAGCGCTACAAAAACAGC	2682		
Db	3334	GATCGGGCGATATCGGCTATCTCAAGGCGCTGTTCTCTACGAGCGCTACAAAAACAGC	3393		
Qy	2683	ATCAGCGCAGCAGCGGTGCGGAGCAACATGCGGAAGGCGAGCTCAACGCGACGCTGATG	2742		
Db	3394	ATCAGCGCAGCAGCGGTGCGGAGCAACATGCGGAAGGCGAGCTCAACGCGACGCTGATG	3453		
Qy	2743	CAGTGGGCGCACTGGGCGGTGTCAACGTTTCCGTCGCGCAACGGGAGATTTGACGGTC	2802		
Db	3454	CAGTGGGCGCACTGGGCGGTGTCAACGTTTCCGTCGCGCAACGGGAGATTTGACGGTC	3513		
Qy	2803	GAAGGCGGTCTGGCTACGACCTGCTCAAAACAGGATGATTCGCCGCAAAAGGACGAGTCT	2862		
Db	3514	GAAGGCGGTCTGGCTACGACCTGCTCAAAACAGGATGATTCGCCGCAAAAGGACGAGTCT	3573		
Qy	2863	TTGGGTGGAGCGGCAACAGCCCTACATGGAAGGCACACTGGTTCGAGACTCGCGGGTCTGAAG	2922		
Db	3574	TTGGGTGGAGCGGCAACAGCCCTACATGGAAGGCACACTGGTTCGAGACTCGCGGGTCTGAAG	3633		

QY 2923 CTGTGCGCAACCCCTTGAGCGGATTAAGCCGCTCCTGTTTGTGCAACGGCGGCGGTGGAACCGCGAC 2982  
Db 3634 CTGTGCGCAACCCCTTGAGCGGATTAAGCCGCTCCTGTTTGTCAACGGCGGCGGTGGAACCGCGAC 3693  
QY 2983 CTGAACGGGACGGACTACACGCTTAACGGGCGGCTTTACCGGGCGGACTTCACGACACCGCGC 3042  
Db 3694 CTGAACGGGACGGACTACACGCTTAACGGGCGGCTTTACCGGGCGGACTTCACGACACCGCGC 3753  
QY 3043 AAGACGGGGGACGCAATATGCGGCACACCCCGCTGTTTGGCGGCTGTTGGCGCGGATGTC 3102  
Db 3754 AAGACGGGGGACGCAATATGCGGCACACCCCGCTGTTTGGCGGCTGTTGGCGCGGATGTC 3813  
QY 3103 GAATTCGGCAACGGCTTGAACGGCTTGGACGCTTACAGCTACGCCCGGTTTCCAAACAGTAC 3162  
Db 3814 GAATTCGGCAACGGCTTGAACGGCTTGGACGCTTACAGCTACGCCCGGTTTCCAAACAGTAC 3873  
QY 3163 GGCACACGACGGGACGAGTCGGCGTAGCTACCGGTTTC 3201  
Db 3874 GGCACACGACGGGACGAGTCGGCGTAGCTACCGGTTTC 3912  
RESULT 12  
AX236463  
LOCUS AX236463 4170 bp DNA linear PAT 26-SEP-2001  
DEFINITION Sequence 156 from Patent W00164922.  
ACCESSION AX236463  
VERSION AX236463.1 GI:15796053  
KEYWORDS synthetic construct.  
ORGANISM synthetic construct.  
REFERENCE 1 (bases 1 to 4170)  
AUTHORS Arico,M.B., Comanducci,M.C., Galeotti,C., Masignani,V.C.,  
Giuliani,M.M. and Pizzi,M.C.  
TITLE Heterologous expression of neisserial proteins  
JOURNAL Patent: WO 0164922-A 156 07-SEP-2001;  
Chiron Spa (IT)  
FEATURES  
Location/Qualifiers  
source  
1. 4170  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="961c-983"  
BASE COUNT 1185 a 1138 c 1119 g 728 t  
ORIGIN  
Query Match 25.9%; Score 831; DB 6; Length 4170;  
Best Local Similarity 99.4%; Pred No. 0;  
Matches 1231; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 1963 CTGTGGCTTCCTCGACAGCTCGAAGGCGGCGGAGTGAAGGCGACACGCTGTGC 2022  
Db 2905 CTGTGGCTTCCTCGACAGCTCGAAGGCGGCGGAGTGAAGGCGACACGCTGTGC 2964  
QY 2023 TATTATGCTCGGCGCAATCGGCGACGGACTGCTTGGGAGCGGCACATTCGGGGCC 2082  
Db 2965 TATTATGCTCGGCGCAATCGGCGACGGACTGCTTGGGAGCGGCACATTCGGGGCC 3024  
QY 2083 GCGGCTCTCAACACGCGCTAGAACGGGCGGCAATCTGSAACCTGATGTCGAA 2142  
Db 3025 GCGGCTCTCAACACGCGCTAGAACGGGCGGCAATCTGSAACCTGATGTCGAA 3084  
QY 2143 CTGGATGCTCGGGAATCATCCGCAACACCGGAGAGGTTGAACTCGGCGCCGCCACCGC 2202  
Db 3085 CTGGATGCTCGGGAATCATCCGCAACACCGGAGAGGTTGAACTCGGCGCCGCCACCGC 3144  
QY 2203 ACAGATATCGGGGATCGCCCGCTACGCGGCACTTCCCGGCGAGCGGCACCGGTACAG 2262  
Db 3145 ACAGATATCGGGGATCGCCCGCTACGCGGCACTTCCCGGCGAGCGGCACCGGTACAG 3204  
QY 2263 CATGGGAATGCGCGGACGGTGTACGATCTTCAACAGTCTCGCGCTACCGCTATGCC 2322  
Db 3205 CATGGGAATGCGCGGACGGTGTACGATCTTCAACAGTCTCGCGCTACCGCTATGCC 3264

QY 2323 GACAGTACCGCGCCCATGCGGATATGCAAGGACCGCGCTGAAAGCCGTATCGACGGG 2382  
Db 3265 GACAGTACCGCGCCCATGCGGATATGCAAGGACCGCGCTGAAAGCCGTATCGACGGG 3324  
QY 2383 TTGGACCACAAGCTTACGGGCTGCGGCTATCGCGCAAAACCCACAGACGGTGAAGC 2442  
Db 3325 TTGGACCACAAGCGGACCGGCTGCGGCTATCGCGCAAAACCCACAGACGGTGAAGC 3384  
QY 2443 TGGGAACAGGGGGTGTGAAGGCAAAATGCGCGCAGTACCACAAACCGTGGCATTTGCC 2502  
Db 3385 TGGGAACAGGGGGTGTGAAGGCAAAATGCGCGCAGTACCACAAACCGTGGCATTTGCC 3444  
QY 2503 GCGAAACCGCGGAAATATACAGACAGCGCCACACTGGGCATGGGACACACATGG 2562  
Db 3445 GCGAAACCGCGGAAATATACAGACAGCGCCACACTGGGCATGGGACACACATGG 3504  
QY 2563 AGCGAAACAGTGAANTGCAAAACCCGACACATTTAGTCTGTTTGCAGGCATACGGAC 2622  
Db 3505 AGCGAAACAGTGAANTGCAAAACCCGACACATTTAGTCTGTTTGCAGGCATACGGAC 3564  
QY 2623 GATCGGGCGATATCGGCTATCTCAAGGCGCTTGTCTCTACGGACGCTACAAAACAGC 2682  
Db 3565 GATCGGGCGATATCGGCTATCTCAAGGCGCTTGTCTCTACGGACGCTACAAAACAGC 3624  
QY 2683 ATCAGCCGACAGCAGCGTGGGAGCAACATGCGGAGCGAGCGTCAACGGCACGCTGATG 2742  
Db 3625 ATCAGCCGACAGCAGCGTGGGAGCAACATGCGGAGCGAGCGTCAACGGCACGCTGATG 3684  
QY 2743 CAGCTTGGGCGCACTGGGCGGTGTCACGCTTCGTTTGGCGAACGGGAGATTGACGGTC 2802  
Db 3685 CAGCTTGGGCGCACTGGGCGGTGTCACGCTTCGTTTGGCGAACGGGAGATTGACGGTC 3744  
QY 2803 GAAGCGGCTCTCGCTACGACCTGCTCAAAACAGGATGATTCGCCGCAAAAGCGAGTGT 2862  
Db 3745 GAAGCGGCTCTCGCTACGACCTGCTCAAAACAGGATGATTCGCCGCAAAAGCGAGTGT 3804  
QY 2863 TTGGCTGAGGCGGCAACAGCCTCCTGAAGGCACACTGTGTCGGACTCGCGGGTCTGAAG 2922  
Db 3805 TTGGCTGAGGCGGCAACAGCCTCCTGAAGGCACACTGTGTCGGACTCGCGGGTCTGAAG 3864  
QY 2923 CTGTGCGCAACCTTGAAGGATTAAGCCGCTCCTGTTTGAACGGGCGGCGTGAACCGCGAC 2982  
Db 3865 CTGTGCGCAACCTTGAAGGATTAAGCCGCTCCTGTTTGAACGGGCGGCGTGAACCGCGAC 3924  
QY 2983 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGGACTTCGACGCAACCGCG 3042  
Db 3925 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGGACTTCGACGCAACCGCG 3984  
QY 3043 AAGACGGGGGACGCAATATGCGGCACACCCCGCTGTTTGGCGGCTGCGGCGGATGTC 3102  
Db 3985 AAGACGGGGGACGCAATATGCGGCACACCCCGCTGTTTGGCGGCTGCGGCGGATGTC 4044  
QY 3103 GAATTCGCAACGGGCTGGAACGGCTTGGCACGTTTACAGCTACCGCGGTTTCCAAACAGTAC 3162  
Db 4045 GAATTCGCAACGGGCTGGAACGGCTTGGCACGTTTACAGCTACCGCGGTTTCCAAACAGTAC 4104  
QY 3163 GGCACACGACGGGACGAGTCGGCGTAGCTACCGGTTTC 3201  
Db 4105 GGCACACGACGGGACGAGTCGGCGTAGCTACCGGTTTC 4143

RESULT 13  
AX239749  
LOCUS AX239749 4170 bp DNA linear PAT 26-SEP-2001  
DEFINITION Sequence 51 from Patent W00164920.  
ACCESSION AX239749  
VERSION AX239749.1 GI:15797385  
KEYWORDS synthetic construct.  
ORGANISM synthetic construct.  
REFERENCE 1 (bases 1 to 4170)  
AUTHORS Arico,M.B., Comanducci,M., Galeotti,C., Masignani,V., Giuliani,M.M.

TITLE  
Hybrid expression of neisserial proteins  
Patent: WO 0164920-A 51 07-SEP-2001;  
Chiron Spa (IT)

FEATURES  
Location/Qualifiers

Source  
I. .4179  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="961c-983"

BASE COUNT 1185 a 1138 c 1119 g 728 t

ORIGIN  
Query Match 25.9%; Score 831; DB 6; Length 4170;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 1231; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1963 CTGCTGGCTTCCCTCGACAGCGTCGAAAAACAGCGGGGAGTGAAGGCGACACGCTGTGC 2022  
Db 1963 TTTGCTGGCTTCCCTCGACAGCGTCGAAAAACAGCGGGGAGTGAAGGCGACACGCTGTGC 2964  
QY 2023 TATTATGTCCGTCGCGCAATGCGGACGACTGCTTCCGACGCGGCACATTCGCGGCC 2082  
Db 2023 TATTATGTCCGTCGCGCAATGCGGACGACTGCTTCCGACGCGGCACATTCGCGGCC 3024  
QY 2083 GCCGGTCTGAACACACGCCGCTAGAACAGGCGGCGAGCAATCTGGAACACCTGTATGTCGAA 2142  
Db 3025 GCCGGTCTGAACACACGCCGCTAGAACAGGCGGCGAGCAATCTGGAACACCTGTATGTCGAA 3084  
QY 2143 CTGGATGCTCCGAATCATCCGCAACACCGGACAGCGTGAACACTGGGCGCGCGACCGC 2202  
Db 3085 CTGGATGCTCCGAATCATCCGCAACACCGGACAGCGTGAACACTGGGCGCGCGACCGC 3144  
QY 2203 ACAGATATGCGGGGATCGCCCTTACGGCGCAACTTTCCGCGCAGCGGCGAGCGGTACAG 2262  
Db 3145 ACAGATATGCGGGGATCGCCCTTACGGCGCAACTTTCCGCGCAGCGGCGAGCGGTACAG 3204  
QY 2263 CATGCGAATGCCCGGACGCGTGTACGATCTTCAACAGTCTCCCGCTACCGCTCTATGCC 2322  
Db 3205 CATGCGAATGCCCGGACGCGTGTACGATCTTCAACAGTCTCCCGCTACCGCTCTATGCC 3264  
QY 2323 GACAGTACCGCCCGCCATGCCGATATGCGAGGACGCGGCTGAAAGCGGTATCGGACGGG 2382  
Db 3265 GACAGTACCGCCCGCCATGCCGATATGCGAGGACGCGGCTGAAAGCGGTATCGGACGGG 3324  
QY 2383 TTGGACCAACACCTTACGGGTCTGCGGTCATCGCGCAACCAACAGGAGCGTGAACG 2442  
Db 3325 TTGGACCAACACGCGACGGGTCTGCGGTCATCGCGCAACCAACAGGAGCGTGAACG 3384  
QY 2443 TGGGAACAGGGCGGTGTGAAGGCAAAATGCGCGGAGTACCCAAACCGCTCGGCATTGCC 2502  
Db 3385 TGGGAACAGGGCGGTGTGAAGGCAAAATGCGCGGAGTACCCAAACCGCTCGGCATTGCC 3444  
QY 2503 GCGAAACCGCGCAAAATACGACAGCAGCGCGGCACACTGGGCATGGGACACACGACATGG 2562  
Db 3445 GCGAAACCGCGCAAAATACGACAGCAGCGCGGCACACTGGGCATGGGACGACGACATGG 3504  
QY 2563 AGCGAAACAGTGCAAAATGCAAAACCGCAGCATTTAGTCTGTTTCAGGCATACGGGCAC 2622  
Db 3505 AGCGAAACAGTGCAAAATGCAAAACCGCAGCATTTAGTCTGTTTCAGGCATACGGGCAC 3564  
QY 2623 GATGCGGGCGATATCGGCTATCTCAAAGGCGCTGTCTCTACGAGCGCTACAAAAACAGC 2682  
Db 3565 GATGCGGGCGATATCGGCTATCTCAAAGGCGCTGTCTCTACGAGCGCTACAAAAACAGC 3624  
QY 2683 ATCAGCGGACGACCGGTGCGGACGAAATGCGGAAGGAGCGGTCAACGCGACGCTGTATG 2742  
Db 3625 ATCAGCGGACGACCGGTGCGGACGAAATGCGGAAGGAGCGGTCAACGCGACGCTGTATG 3684  
QY 2743 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTCCGCGACGCGAGATTTGACGGTTC 2802  
Db 3685 CAGCTGGGCGCACTGGGCGGTGTCAACGTTTCCGTTTCCGCGACGCGAGATTTGACGGTTC 3744  
QY 2803 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCTTCCGCCGCAAAAGGACGCTGCT 2862

Db 3745 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCTATTCGCCGAAAAAGGACGTGCT 3804  
QY 2863 TTGGCTGGAGCGGCAACAGCCTCCTACTGAAGGACACACTGTTCGGACTTCGGGGTCTGAAG 2922  
Db 3805 TTGGCTGGAGCGGCAACAGCCTCCTACTGAAGGACAGCGTGTTCGGACTTCGGGGTCTGAAG 3864  
QY 2923 CTGTCGCAACCCCTTGAGCGATAAAGCGCTCTGTTTTCGCAACGCGGCGGTGGAACGCGAC 2982  
Db 3865 CTGTCGCAACCCCTTGAGCGATAAAGCGCTCTGTTTTCGCAACGCGGCGGTGGAACGCGAC 3924  
QY 2983 CTGAACGCGACGCGACTACACGTAACGCGGCGGCTTTTACCAGCGCGACTGCAGCAACCGGC 3042  
Db 3925 CTGAACGCGACGCGACTACACGTAACGCGGCGGCTTTTACCAGCGCGACTGCAGCAACCGGC 3984  
QY 3043 AACACGCGGCGACGCAATATGCGCGCACACCGCGCTGTTCCCGGTCTGGCGCGGATGTC 3102  
Db 3985 AACACGCGGCGACGCAATATGCGCGCACACCGCGCTGTTCCCGGTCTGGCGCGGATGTC 4044  
QY 3103 GAATTTCGGCAACCGCTTGGAAACGCTTGGCACGTTACAGCTACGCGGTTCCAAACAGTAC 3162  
Db 4045 GAATTTCGGCAACCGCTTGGAAACGCTTGGCACGTTACAGCTACGCGGTTCCAAACAGTAC 4104  
QY 3163 GGCAACACAGCGGACGAGTCCGCGTAGGCTACCGGTTTC 3201  
Db 4105 GGCAACACAGCGGACGAGTCCGCGTAGGCTACCGGTTTC 4143  
RESULT 14  
AX236413  
LOCUS AX236413 4179 bp DNA linear PAT 26-SEP-2001  
DEFINITION Sequence 106 from Patent WO016492..  
ACCESSION AX236413  
VERSION AX236413.1 GI:15796029  
KEYWORDS synthetic construct.  
SOURCE synthetic construct.  
ORGANISM artificial sequences.  
REFERENCE 1 (bases 1 to 4179)  
AUTHORS Arico,M.B., Comanducci,M.C., Galeotti,C.C., Massignani,V.C.,  
Guiliani,M.M. and Pizzi,M.C.  
TITLE Heterologous expression of neisserial proteins  
JOURNAL Patent: WO 0164922-A 106 07-SEP-2001;  
Chiron Spa (IT)  
FEATURES  
Source  
I. .4179  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="deltaG983-961c"  
BASE COUNT 1187 a 1140 c 1121 g 731 t  
ORIGIN  
Query Match 25.9%; Score 831; DB 6; Length 4179;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 1231; Conservative 0; Mismatches 8; Indels 0; Gaps 0;



|||||  
Db 2146 ACAGATATGCCGGGATCGCCCTACGGCGCAACTTCCGGCGCAGCGGACGCCGTACAG 2205  
Qy 2263 CATGGGAATGCCCGCGACGGGTACGCATCTTCAACAGTCTCGCGCTACCGCTCTATGCC 2322  
Db 2206 CATGGGAATGCCCGCGACGGGTACGCATCTTCAACAGTCTCGCGCTACCGCTCTATGCC 2265  
Qy 2323 GACGTACCGCGCGCCATGCGCATATGCGAGGACCGCGGCTGAAAGCCGTATCGACGGG 2382  
Db 2266 GACAGTACCGCGCGCCATGCGCATATGCGAGGACCGCGCTGAAAGCCGTATCGACGGG 2325  
Qy 2383 TTGGACCACACGCTACGGGTCTGGCGTCTATCGCGCAACCCACAGACGGGTGAAGC 2442  
Db 2326 TTGGACCACACGCGACGGGTCTGGCGGTCTATCGCGCAACCCACAGACGGGTGAAGC 2385  
Qy 2443 TGGGAACAGGGCGGTGTTCAAGCGCAAAATGCGCGCAGTACCCAAACCTCGCGCATTTGCC 2502  
Db 2386 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGAGTACCCAAACCTCGCGCATTTGCC 2445  
Qy 2503 GCGAAACCGCGGAAATACGACAGCGCGCCACACTGGGATGGGACACAGCACATGG 2562  
Db 2446 GCGAAACCGCGGAAATACGACAGCGCGCCACACTGGGATGGGACGACACATGG 2505  
Qy 2563 ACGGAAACAGTGCRAATGCAAAACCGGACAGCATTTAGTCTGTTGCGAGGATACGGCAC 2622  
Db 2506 ACGGAAACAGTGCRAATGCAAAACCGGACAGCATTTAGTCTGTTGCGAGGATACGGCAC 2565  
Qy 2623 GATGGGGGATATCGGCTATCTCAAGGCGCTTCTCTCTACGGAGCTTACAAAACAGC 2682  
Db 2566 GATGGGGGATATCGGCTATCTCAAGGCGCTTCTCTCTACGGAGCTTACAAAACAGC 2625  
Qy 2683 ATCAGCGCGAGCACCGGTGCGGACGAACATGCGGAGGAGCGTCAACGGCACGCTGATG 2742  
Db 2626 ATCAGCGCGAGCACCGGTGCGGACGAACATGCGGAGGAGCGTCAACGGCACGCTGATG 2685  
Qy 2743 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGGCGCAACGGGAGATTTGACGGTC 2802  
Db 2686 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGGCGCAACGGGAGATTTGACGGTC 2745  
Qy 2803 GAAGGCGGTCTCGGCTACGACCTGCTCAACAGAGTATGTCCTCCGCAAAAGGCACTGCT 2862  
Db 2746 GAAGGCGGTCTCGGCTACGACCTGCTCAACAGAGTATGTCCTCCGCAAAAGGCACTGCT 2805  
Qy 2863 TTGGGCTGAGCGGCAACAGCTCACTAAGGCACACTGGTCCGACTCGCGGTCTGAAG 2922  
Db 2806 TTGGGCTGAGCGGCAACAGCTCACTAAGGCACACTGGTCCGACTCGCGGTCTGAAG 2865  
Qy 2923 CTGTCCCAACCTTGAGCGATTAAGCCGTCTCTTTTGCAACGGCGGGGTGGAACCGCAC 2982  
Db 2866 CTGTCCCAACCTTGAGCGATTAAGCCGTCTCTTTTGCAACGGCGGGGTGGAACCGCAC 2925  
Qy 2983 CTGAACGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGGACTTGCAGCAACCGGC 3042  
Db 2926 CTGAACGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGGACTTGCAGCAACCGGC 2985  
Qy 3043 AAGACGGGGCACGCAATATGCCGACACCCCGCTGTTGTCGGGCTGCGGCGGATGTC 3102  
Db 2986 AAGACGGGGCACGCAATATGCCGACACCCCGCTGTTGTCGGGCTGCGGCGGATGTC 3045  
Qy 3103 GAATTCGGCAACGGCTTGGAAACGGCTTGGCAAGCTTACAGCTACCGCGGTTCCAAACAGTAC 3162  
Db 3046 GAATTCGGCAACGGCTTGGAAACGGCTTGGCAAGCTTACAGCTACCGCGGTTCCAAACAGTAC 3105  
Qy 3163 GCGAACACAGCGGAGGAGTTCGGGCTAGCGCTACCGGTTCC 3201  
Db 3106 GCGAACACAGCGGAGGAGTTCGGGCTAGCGCTACCGGTTCC 3144

RESULT 15  
AX239721  
LOCUS AX239721 4179 bp DNA linear PAT 26-SEP-2001  
DEFINITION Sequence 23 from Patent WO0164920.  
ACCESSION AX239721

AX239721.1 GI:15797372  
synthetic construct.  
synthetic construct  
artificial sequences.  
1 (bases 1 to 4179)  
Arico,M.B., Comanducci,M., Galeotti,C., Masignani,V., Giuliani,M.M.  
and Pizza,M.  
Hybrid expression of neisserial proteins  
Patent: WO 0164920-A 23 07-SEP-2001;  
Chiron Spa (IT)  
Location/Qualifiers  
1. 4179  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="deltaG983-961c"  
BASE COUNT 1187 a 1140 c 1121 g 731 t  
ORIGIN  
Query Match 25.9%; Score 831; DB 6; Length 4179;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 1231; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
Qy 1963 CTGCTGGCTTCCTCGACAGCTCGGAAACACAGCGGGCAGTGAAGGCGACACGCTGTCC 2022  
Db 1906 CTGCTGGCTTCCTCGACAGCTCGGAAACACAGCGGGCAGTGAAGGCGACACGCTGTCC 1965  
Qy 2023 TATTATGTCGTCGCGCAATCGGCACGACTCTCGGCAGCGCAGCATTTCCCGGCC 2082  
Db 1966 TATTATGTCGTCGCGCAATCGGCACGACTCTCGGCAGCGCAGCATTTCCCGGCC 2025  
Qy 2083 GCCGCTCTGAAACACGCGGTAGAACAGGCGCGCAGCAATCTGGAACCTGATGTCGAA 2142  
Db 2026 GCCGCTCTGAAACACGCGGTAGAACAGGCGCGCAGCAATCTGGAACCTGATGTCGAA 2085  
Qy 2143 CTGGATGCTTCGAATCATCCGCAACACACCGGAGAGCGTTGAAACTTCGCGCCGCCACCGC 2202  
Db 2086 CTGGATGCTTCGAATCATCCGCAACACACCGGAGAGCGTTGAAACTTCGCGCCGCCACCGC 2145  
Qy 2203 ACAGATATGCGCGGATCGCGCCCTACGCGCAACTTTCCGCGCAGCGCGCAGCGGTACAG 2262  
Db 2146 ACAGATATGCGCGGATCGCGCCCTACGCGCAACTTTCCGCGCAGCGCGCAGCGGTACAG 2205  
Qy 2263 CATGCAATGCGCGCGAGGTGTACGCATCTTCAACAGTCTCGCGCTACCGCTCTATGCC 2322  
Db 2206 CATGCAATGCGCGCGAGCGTGTACGCATCTTCAACAGTCTCGCGCTACCGCTCTATGCC 2265  
Qy 2323 GACGTACCGCGCCCATGCCGATATGAGGACCGCGGCTGAAAGCCGTATCGGACGGG 2382  
Db 2266 GACGTACCGCGCCCATGCCGATATGAGGACCGCGGCTGAAAGCCGTATCGGACGGG 2325  
Qy 2383 TTGGACCACACGCTACGGGTCTGCGGCTATCGCGCAACCCACAGACGGGTGAAGC 2442  
Db 2326 TTGGACCACACGCGACGGGTCTGCGGCTATCGCGCAACCCACAGACGGGTGAAGC 2385  
Qy 2443 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGCAGTACCCAAACCTCGCGCATTTGCC 2502  
Db 2386 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGCAGTACCCAAACCTCGCGCATTTGCC 2445  
Qy 2503 GCGAAACCGGGGAAATACGACAGCGCGCCACACTGGGATGGGACACAGCACATGG 2562  
Db 2446 GCGAAACCGGGGAAATACGACAGCGCGCCACACTGGGATGGGACGACACATGG 2505  
Qy 2563 ACGGAAACAGTGCRAATGCAAAACCGGACAGCATTTAGTCTGTTGCGAGGATACGGCAC 2622  
Db 2506 ACGGAAACAGTGCRAATGCAAAACCGGACAGCATTTAGTCTGTTGCGAGGATACGGCAC 2565  
Qy 2623 GATGGGGGATATCGGCTATCTCAAGGCGCTTCTCTCTACGGAGCTTACAAAACAGC 2682  
Db 2566 GATGGGGGATATCGGCTATCTCAAGGCGCTTCTCTCTACGGAGCTTACAAAACAGC 2625  
Qy 2683 ATCAGCGCGAGCACCGGTGCGGACGAACATGCGGAGGAGCGTCAACGGCACGCTGATG 2742  
|||||



Db 2626 ATCAGCCGACGACCCGGTGGGAGCAATGCGGAAGGCGAGTCAACAGCGACGCTGATG 2685  
QY 2743 CAGCTGGCGCACTGGCGGTGTCAACGTTCCGTTCCGCAACAGCGGAGATTTGACGGTC 2802  
Db 2686 CAGCTGGCGCACTGGCGGTGTCAAGTTCCGTTTCGCGCAACAGCGGAGATTTGACGGTC 2745  
QY 2803 GAAGGGGTCTGGCTACGACCTGCTCAACAGAGATGATTCGCCGCAAAAGGAGTGCT 2862  
Db 2746 GAAGGGGTCTGGCTACGACCTGCTCAACAGAGATGATTCGCCGCAAAAGGAGTGCT 2805  
QY 2863 TTGGGTGGAGCGCAACAGCGCTCACTGAAGGCACACTGTCGACGTCGCGGTCTGAAG 2922  
Db 2806 TTGGGTGGAGCGCAACAGCGCTCACTGAAGGCACACTGTCGACGTCGCGGTCTGAAG 2865  
QY 2923 CTGTGCAACCCCTTGACGATAAAGCGTCTCTGTTTCAACAGCGCGCGCTGGAACGCGAC 2982  
Db 2866 CTGTGCAACCCCTTGACGATAAAGCGTCTCTGTTTCAACAGCGCGCGCTGGAACGCGAC 2925  
QY 2983 CTGAACGGACGCGACTACAGCGTAACGGGCGGTCTTACCGCGCGACTGCGAGCAACCGGC 3042  
Db 2926 CTGAACGGACGCGACTACAGCGTAACGGGCGGTCTTACCGCGCGACTGCGAGCAACCGGC 2985  
QY 3043 AAGACGGGCGACGCAATATGCGCGCACACCGCTGTTGCGGTCTGCGGCGGATGTC 3102  
Db 2986 AAGACGGGCGACGCAATATGCGCGCACACCGCTGTTGCGGTCTGCGGCGGATGTC 3045  
QY 3103 GAATTCGGCAACCGCTGGAACGGCTTGGCACGTTACAGCTACGCGCGGTTCCTCAACAGTAC 3162  
Db 3046 GAATTCGGCAACCGCTGGAACGGCTTGGCACGTTACAGCTACGCGCGGTTCCTCAACAGTAC 3105  
QY 3163 GGAACACAGCGGACGAGTGGCGTATAGGTACCGGTC 3201  
Db 3106 GGAACACAGCGGACGAGTGGCGTATAGGTACCGGTC 3144

RESULT 16  
AX236457  
LOCUS AX236457 4335 bp DNA linear PAT 26-SEP-2001  
DEFINITION Sequence 150 from Patent WO0164922.  
ACCESSION AX236457  
VERSION AX236457.1 GI:15796050  
KEYWORDS  
SOURCE synthetic construct.  
ORGANISM synthetic construct.  
REFERENCE 1 (bases 1 to 4335)  
AUTHORS Arico,M.B., Comanducci,M.C., Galeotti,C.C., Massignani,V.C.,  
Guillani,M. and Pizzi,M.C.  
TITLE Heterologous expression of neisserial proteins  
JOURNAL Patent: WO 0164922-A 150 07-SEP-2001;  
Chiron Spa (IT)  
FEATURES  
Source  
1..4335  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="961-983"

BASE COUNT 1217 a 1189 c 1165 g 764 t  
ORIGIN

Query Match 25.9%; Score 831; DB 6; Length 4335;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 1231; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1963 CTGCTGGCTTCCCTCGACACGCTGCAAAACACAGCGGCGAGTGAAGCGACACGCTGTCC 2022  
Db 3070 CTGCTGGCTTCCCTCGACACGCTGCAAAACACAGCGGCGAGTGAAGCGACACGCTGTCC 3129  
QY 2023 TATTATGTCGTCGCGCAATGGCGACGACTGCTTCGCGACGCGGCACATTCGCGGCC 2082  
Db 3130 TATTATGTCGTCGCGCAATGGCGACGACTGCTTCGCGACGCGGCACATTCGCGGCC 3189  
QY 2083 GCCGGTCTGAACACGCGGTAGAACAGGCGCGAGCAATCTGGAACACCTGATGGTCGAA 2142  
|||||

Db 3190 GCCGGTCTGAACACACGCGGTAGAACAGGCGGCGAGCAATCTGGAACACCTGATGGTCGA 3249  
QY 2143 CTGATGCCCTCCGAATCATCCGCAACACCGAGACGCTTGAAACTCGCGCGCGACCGC 2202  
Db 3250 CTGATGCCCTCCGAATCATCCGCAACACCGAGACGCTTGAAACTCGCGCGCGACCGC 3309  
QY 2203 ACAGATATGCCGGGCATCCGCCCTACGGGCGCAACTTTTCGCGCGAGCGGCGAGCTACAG 2262  
Db 3310 ACAGATATGCCGGGCATCCGCCCTACGGGCGCAACTTTTCGCGCGAGCGGCGAGCTACAG 3369  
QY 2263 CATGCGAATGCCGCCGCGAGGTGTAGCATCTTCAACAGTCTCGCGCGCTACGCTCTATGCC 2322  
Db 3370 CATGCGAATGCCGCCGCGAGGTGTAGCATCTTCAACAGTCTCGCGCGCTACGCTCTATGCC 3429  
QY 2323 GACAGTACCGCCGCCATGCCGATATCGAGGACCGCGCTGAAAGCCGATTCGAGCGG 2382  
Db 3430 GACAGTACCGCCGCCATGCCGATATCGAGGACCGCGCTTGAAGCCGATTCGAGCGG 3489  
QY 2383 TTGSACACACACGCTACGGGTCTGCGGTCTATCGCGCAAAACCCCAACAGGACGCTGGAACG 2442  
Db 3490 TTGSACACACACGCTACGGGTCTGCGGTCTATCGCGCAAAACCCCAACAGGACGCTGGAACG 3549  
QY 2443 TGGGAACAGGGCGGTGTGAAGGCAAAATGCGCGGCGAGTACCCCAACCGCTCGGCATTCG 2502  
Db 3550 TGGGAACAGGGCGGTGTGAAGGCAAAATGCGCGGCGAGTACCCCAACCGCTCGGCATTCG 3609  
QY 2503 GCGAAACCGCGCAAAATACGACAGCGCGCACACTGGGCGTGGGACACAGCACATGG 2562  
Db 3610 GCGAAACCGCGCAAAATACGACAGCGCGCACACTGGGCGTGGGACACAGCACATGG 3669  
QY 2563 AGCGAAACAGTGCAAATGCAAAACCGACAGCATTAAGTCTGTTTTCGAGGCGATACGCGAC 2622  
Db 3670 AGCGAAACAGTGCAAATGCAAAACCGACAGCATTAAGTCTGTTTTCGAGGCGATACGCGAC 3729  
QY 2623 GATGCGGGCGATATCGGTCTATCTCAAAAGGCTGTTCTCTACGACGCTTCAAAAACAGC 2682  
Db 3730 GATGCGGGCGATATCGGTCTATCTCAAAAGGCTGTTCTCTACGACGCTTCAAAAACAGC 3789  
QY 2683 ATCAGCGCGACGACCGGTGCGGAGCAATCGGAAGCGAGCTCAACGACGCGCTCATG 2742  
Db 3790 ATCAGCGCGACGACCGGTGCGGAGCAATCGGAAGCGAGCTCAACGACGCGCTCATG 3849  
QY 2743 CAGCTGGCGCACTGGCGGTGTCAACGTTCCGTTTCGCGCAACGGGAGATTTGACGGTC 2802  
Db 3850 CAGCTGGCGCACTGGCGGTGTCAACGTTCCGTTTCGCGCAACGGGAGATTTGACGGTC 3909  
QY 2803 GAAGGGGTCTGGCTAGCAGCTGCTCAACAGGATGATTCGCCGCAAAAGGAGCTGCT 2862  
Db 3910 GAAGGGGTCTGGCTAGCAGCTGCTCAACAGGATGATTCGCCGCAAAAGGAGCTGCT 3969  
QY 2863 TTGGGTGGAGCGCAACAGCTTCACTGAAGGCACACTGGTTCGAGCTCGCGGTCTGAAG 2922  
Db 3970 TTGGGTGGAGCGCAACAGCTTCACTGAAGGCACACTGGTTCGAGCTCGCGGTCTGAAG 4029  
QY 2923 CTGTGCAACCCCTTGAGCGATAAAGCGTCTCTGTTTCAACAGCGCGCGCTGGAACGCGAC 2982  
Db 4030 CTGTGCAACCCCTTGAGCGATAAAGCGTCTCTGTTTCAACAGCGCGCGCTGGAACGCGAC 4089  
QY 2983 CTGAACGGACGCGACTACAGGTAACGGGCGGCTTTTACCGCGCGACTGAGCAACCGGC 3042  
Db 4090 CTGAACGGACGCGACTACAGGTAACGGGCGGCTTTTACCGCGCGACTGAGCAACCGGC 4149  
QY 3043 AAGACGGGCGACGCAATATGCGCGCACACCGCTGCTGCGGTCTGGGCGGAGTGC 3102  
Db 4150 AAGACGGGCGACGCAATATGCGCGCACACCGCTGCTGCGGTCTGGGCGGAGTGC 4209  
QY 3103 GAATTCGGCAACCGCTGGAACGGCTTGGCACGTTTACAGCTACGCGGTTCCTCAACAGTAC 3162  
Db 4210 GAATTCGGCAACCGCTGGAACGGCTTGGCACGTTTACAGCTACGCGGTTCCTCAACAGTAC 4269  
QY 3163 GGAACACAGCGGACGAGTGGCGTATAGGTACCGGTC 3201  
Db 4270 GGAACACAGCGGACGAGTGGCGTATAGGTACCGGTC 4308  
|||||

RESULT 17	AX239743	AX239743	4335 bp	DNA	linear	PAT 26-SEP-2001
LOCUS	Sequence 45 from Patent WO0164920.					
DEFINITION	AX239743					
ACCESSION	AX239743.1	GI:15797382				
VERSION						
KEYWORDS	synthetic construct.					
SOURCE	synthetic construct					
ORGANISM	artificial sequences.					
REFERENCE	1 (bases 1 to 4335)					
AUTHORS	Arlico,M.B., Comanducci,M., Galeotti,C., Masignani,V., Giuliani,M.M. and Pizzo,M.					
TITLE	Hybrid expression of neisserial proteins					
JOURNAL	Patent: WO 0164920-A 45 07-SEP-2001;					
FEATURES	Chiron Spa (IT)					
source	Location/Qualifiers					
	1..4335					
BASE COUNT	1217 a 1189 c 1165 g 764 t					
ORIGIN						
	Query Match 25.9%; Score 831; DB 6; Length 4335;					
	Best Local Similarity 99.4%; Pred. No. 0;					
	Matches 1231; Conservative 0; Mismatches 8; Indels 0; Gaps 0;					
Qy	1963 CTGCTGGCTTCCCTCGACAGCGTCGAAAAACACGGCGCAGTGAAGGCGACACGCTGTGCC	2022				
Db	3070 CTGCTGGCTTCCCTCGACAGCGTCGAAAAACACGGCGCAGTGAAGGCGACACGCTGTGCC	3129				
Qy	2023 TATTATGTCGTCGGCGGAATTCGGCAGCGGACTGCTTCGGCAGCGGCACATTCGCGGCC	2082				
Db	3130 TATTATGTCGTCGGCGGAATTCGGCAGCGGACTGCTTCGGCAGCGGCACATTCGCGGCC	3189				
Qy	2083 GCGGTCTGAAACACGCCGTAGAACAGGGCGGCACAACTCTGGAACACCTGATGTCGAA	2142				
Db	3190 GCGGTCTGAAACACGCCGTAGAACAGGGCGGCACAACTCTGGAACACCTGATGTCGAA	3249				
Qy	2143 CTGGATGCTCCGAAATCATCCGCAACACCCGAGACGGTTGAAACTGCGCGCCGCGACCGC	2202				
Db	3250 CTGGATGCTCCGAAATCATCCGCAACACCCGAGACGGTTGAAACTGCGCGACCGACCGC	3309				
Qy	2203 ACAGATATGCGGGGATCGGCCCTACGGCGCACTTTCCGGCGACGGCAGCGCTACAG	2262				
Db	3310 ACAGATATGCGGGGATCGGCCCTACGGCGCACTTTTCGGCGCAGCGCAGCGCTACAG	3369				
Qy	2263 CATCGGAATGCGCGGAGGTTGACGCACTTTCACAGTCTCGCGCTACCGTCTATGCC	2322				
Db	3370 CATCGGAATGCGCGGAGGTTGACGCACTTTCACAGTCTCGCGCTACCGTCTATGCC	3429				
Qy	2323 GACAGTACCGCGCCCATCGCGATATGAGGAGCAGCCGGCTGAAAGCCGCTATCGGACGGG	2382				
Db	3430 GACAGTACCGCGCCCATCGCGATATGAGGAGCAGCCGGCTGAAAGCCGCTATCGGACGGG	3489				
Qy	2383 TTGGACCAACAGCTTACGGGTCTGCGGTCTATCGCGCAACCCACAGAGCGGTGGAACG	2442				
Db	3490 TTGGACCAACAGCGGAGCGGTCTGCGGTCTATCGCGCAACCCACAGAGCGGTGGAACG	3549				
Qy	2443 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGCAGTACCCAAACCGTCGGCATTGCC	2502				
Db	3550 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGCAGTACCCAAACCGTCGGCATTGCC	3609				
Qy	2503 GCGAAAAACCGGCAAAATACGACAGCAGCGCCACACTTGGGCATGGGACACAGCACATGG	2562				
Db	3610 GCGAAAAACCGGCAAAATACGACAGCAGCGCCACACTTGGGCATGGGACAGCACATGG	3669				
Qy	2563 AGCGAAAAACAGTGAATGCAAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC	2622				
Db	3670 AGCGAAAAACAGTGAATGCAAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC	3729				

QY 2023 TATTATGTCGTCGCGCAATGCGGACGCGACTGCTTCGGCAGCGGCACATTCGCGGCC 2082  
Db 1966 TATTATGTCGTCGCGCAATGCGGACGCGACTGCTTCGGCAGCGGCACATTCGCGGCC 2025  
QY 2083 GCCGGTCTGAACACACCGCTAGAACAGGCGGGCAGCAATCTGGAACACCTGATGGTCGAA 2142  
Db 2026 GCCGGTCTGAACACACCGCTAGAACAGGCGGGCAGCAATCTGGAACACCTGATGGTCGAA 2085  
QY 2143 CTGGATGCTCCGCAATCATCCGCAACACCGAGACAGCGTGTGAACCTGCGGCGCGGACCGC 2202  
Db 2086 CTGGATGCTCCGCAATCATCCGCAACACCGAGACAGCGTGTGAACCTGCGGCGCGGACCGC 2145  
QY 2203 ACAGATATGCGGGCATCCGCCCTTACGGCGCAACTTTTCGGCGACGCGGACCGCTACAG 2262  
Db 2146 ACAGATATGCGGGCATCCGCCCTTACGGCGCAACTTTTCGGCGACGCGGACCGCTACAG 2205  
QY 2263 CATGCCAATGCCCGCCGACCGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2322  
Db 2206 CATGCCAATGCCCGCCGACCGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2265  
QY 2323 GACAGTACCGCCGCCCATGCGGATATGACGGGACGCGGCTGAAGCGCTATCGGACGGG 2382  
Db 2266 GACAGTACCGCCGCCCATGCGGATATGACGGGACGCGGCTGAAGCGCTATCGGACGGG 2325  
QY 2383 TTGGACCAACAGCTACGGGTCTGCGCGTCATCGCGCAACCCACAGGACGGTGGAAAG 2442  
Db 2326 TTGGACCAACAGCTACGGGTCTGCGCGTCATCGCGCAACCCACAGGACGGTGGAAAG 2385  
QY 2443 TGGGAACAGGGCGGTGTGAAGGCAAAATTCGCGGGCAGTACCCAAACCGTTCGGCATTGC 2502  
Db 2386 TGGGAACAGGGCGGTGTGAAGGCAAAATTCGCGGGCAGTACCCAAACCGTTCGGCATTGC 2445  
QY 2503 GCGAAAACCGCGGAAATACGACAGCAGCGCGCCACACTGGGCATGGGACACACACATGG 2562  
Db 2446 GCGAAAACCGCGGAAATACGACAGCAGCGCGCCACACTGGGCATGGGACACACACATGG 2505  
QY 2563 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTTAGTCTGTTTCAGGCATACGGCAC 2622  
Db 2506 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTTAGTCTGTTTCAGGCATACGGCAC 2565  
QY 2623 GATCGGGCGCATATCGGCTATCTCAAGCGCTGTTCTCTACGGACGCTACAAAAACAGC 2682  
Db 2566 GATCGGGCGCATATCGGCTATCTCAAGCGCTGTTCTCTACGGACGCTACAAAAACAGC 2625  
QY 2683 ATCAGCGCAGCACCGCTGCGGACGAACATGCGGAAGGACGCTCAACGACGCTGATG 2742  
Db 2626 ATCAGCGCAGCACCGCTGCGGACGAACATGCGGAAGGACGCTCAACGACGCTGATG 2685  
QY 2743 CAGCTGGGCGCACTGGGCGGTGTCAACGTTTCCGTTTTCGCGCAACGGGAGATTTGACGGTC 2802  
Db 2686 CAGCTGGGCGCACTGGGCGGTGTCAACGTTTCCGTTTTCGCGCAACGGGAGATTTGACGGTC 2745  
QY 2803 GAAGCGGTCTGGCTACGACCTGCTCAACAGGATGCTATGCCGCAAAAGGACATGCT 2862  
Db 2746 GAAGCGGTCTGGCTACGACCTGCTCAACAGGATGCTATGCCGCAAAAGGACATGCT 2805  
QY 2863 TTGGGCTGAGCGGCAACAGCTTCACTGAAGGCACACTGCTCGGACTCGCGGTCTGAAG 2922  
Db 2806 TTGGGCTGAGCGGCAACAGCTTCACTGAAGGCACACTGCTCGGACTCGCGGTCTGAAG 2865  
QY 2923 CTGTGCAACCCCTTGAAGCGATAAAGCGTCTCTTTTGAACGCGGCGCTGGAACCGGAC 2982  
Db 2866 CTGTGCAACCCCTTGAAGCGATAAAGCGTCTCTTTTGAACGCGGCGCTGGAACCGGAC 2925  
QY 2983 CTGAACGGACGCTACAGGTAAAGCGGCTTTTACCGGCGGCTTACGACAAACCGG 3042  
Db 2926 CTGAACGGACGCTACAGGTAAAGCGGCTTTTACCGGCGGCTTACGACAAACCGG 2985  
QY 3043 AAGACGGGGCAGCAATATCGGCACACCGCTGTTGCGGCTTGGCGGGGATGTC 3102  
Db 2986 AAGACGGGGCAGCAATATCGGCACACCGCTGTTGCGGCTTGGCGGGGATGTC 3045

QY 3103 GAATTCGGCAACGGCTGGAAACGGCTTGGCACGTTTACAGCTACGCGGTTCCAAACAGTAC 3162  
Db 3046 GAATTCGGCAACGGCTGGAAACGGCTTGGCACGTTTACAGCTACGCGGTTCCAAACAGTAC 3105  
QY 3163 GCGAACACAGCGGACGAGTGGCGGTAGGTACCGGTTCC 3201  
Db 3106 GCGAACACAGCGGACGAGTGGCGGTAGGTACCGGTTCC 3144  
RESULT 19  
AX239719  
LOCUS AX239719  
DEFINITION Sequence 21 from Patent WO0164920.  
ACCESSION AX239719  
VERSION AX239719.1 GI:15797371  
KEYWORDS  
SOURCE synthetic construct.  
ORGANISM synthetic construct  
artificial sequences.  
REFERENCE 1 (bases 1 to 4344)  
AUTHORS Arico,M.B., Comanducci,M., Galeotti,C., Masignani,V., Giuliani,M.M.  
and Pizza,M.  
TITLE Hybrid expression of neisserial proteins  
JOURNAL Patent: WO 0164920-A 21 07-SEP-2001;  
Chiron Spa (IT)  
FEATURES  
Location/Qualifiers  
1..4344  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="deltaG983-961"  
BASE COUNT 1219 a 1191 c 1167 g 767 t  
ORIGIN  
Query Match 25.9%; Score 831; DB 6; Length 4344;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 1231; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 1963 CTGCTGGCTTCCCTCGACAGCGTCGAAAAACAGCGGCGAGTGAAGCGGACACGCTGTGCC 2022  
Db 1906 CTGCTGGCTTCCCTCGACAGCGTCGAAAAACAGCGGCGAGTGAAGCGGACACGCTGTGCC 1965  
QY 2023 TATTATGTCGTCGCGCAATGCGGACGCGACTGCTTCGGCAGCGGCACATTCGCGGCC 2082  
Db 1966 TATTATGTCGTCGCGCAATGCGGACGCGACTGCTTCGGCAGCGGCACATTCGCGGCC 2025  
QY 2083 GCCGGTCTGAACACACCGCTAGAACAGGCGGCGAGCAATCTGGAACACCTGATGGTCGAA 2142  
Db 2026 GCCGGTCTGAACACACCGCTAGAACAGGCGGCGAGCAATCTGGAACACCTGATGGTCGAA 2085  
QY 2143 CTGGATGCTCCGCAATCATCCGCAACACCGAGACAGCGTGTGAACCTGCGGCGCGGACCGC 2202  
Db 2086 CTGGATGCTCCGCAATCATCCGCAACACCGAGACAGCGTGTGAACCTGCGGCGCGGACCGC 2145  
QY 2203 ACAGATATGCGGGCATCCGCCCTTACGGCGCAACTTTTCGGCGACGCGGACCGCTACAG 2262  
Db 2146 ACAGATATGCGGGCATCCGCCCTTACGGCGCAACTTTTCGGCGACGCGGACCGCTACAG 2205  
QY 2263 CATGCCAATGCCCGCCGACCGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2322  
Db 2206 CATGCCAATGCCCGCCGACCGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2265  
QY 2323 GACAGTACCGCCGCCCATGCGGATATGACGGGACGCGGCTGAAGCGCTATCGGACGGG 2382  
Db 2266 GACAGTACCGCCGCCCATGCGGATATGACGGGACGCGGCTGAAGCGCTATCGGACGGG 2325  
QY 2383 TTGGACCAACAGCTACGGGTCTGCGCGTCATCGCGCAACCCACAGGACGGTGGAAAG 2442  
Db 2326 TTGGACCAACAGCTACGGGTCTGCGCGTCATCGCGCAACCCACAGGACGGTGGAAAG 2385  
QY 2443 TGGGAACAGGGCGGTGTGAAGGCAAAATTCGCGGGCAGTACCCAAACCGTTCGGCATTGC 2502  
Db 2386 TGGGAACAGGGCGGTGTGAAGGCAAAATTCGCGGGCAGTACCCAAACCGTTCGGCATTGC 2445

QY 2503 GCGAAACCGCGGCAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCATGG 2562  
Db 2446 GCGAAACCGCGGCAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCATGG 2505  
QY 2563 AGCGAAACAGTGCATAATGCAAAACCCGACAGCATTAGTCTCTTTGCGAGGCATACGGCAC 2622  
Db 2506 AGCGAAACAGTGCATAATGCAAAACCCGACAGCATTAGTCTCTTTGCGAGGCATACGGCAC 2565  
QY 2623 GATCGGGGGATATCGGCTATCTCAAGGCCCTGTCTCCTACGGAGCGCTACAAAAACAGC 2682  
Db 2566 GATCGGGGGATATCGGCTATCTCAAGGCCCTGTCTCCTACGGAGCGCTACAAAAACAGC 2625  
QY 2683 ATCAGCCGACAGCAGCGGTGCGGACGAACATGCGGAAGCAGCGTCAACGGCACGCTGATG 2742  
Db 2626 ATCAGCCGACAGCAGCGGTGCGGACGAACATGCGGAAGCAGCGTCAACGGCACGCTGATG 2685  
QY 2743 CAGCTGGGGCACTGGCGGGTGTCAACGTTTCGGTTTGGCGCAACGGGAGATTGACGGTTC 2802  
Db 2686 CAGCTGGGGCACTGGCGGGTGTCAACGTTTCGGTTTGGCGCAACGGGAGATTGACGGTTC 2745  
QY 2803 GAAGCGGTCTCGCTACGACCTGCTCAACAGAGTGCATTCCCGGCAAAAGGCAGTGT 2862  
Db 2746 GAAGCGGTCTCGCTACGACCTGCTCAACAGAGTGCATTCCCGGCAAAAGGCAGTGT 2805  
QY 2863 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGCAGCAGTGTGCGGACTCGCGGGTCTGAAG 2922  
Db 2806 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGCAGCAGTGTGCGGACTCGCGGGTCTGAAG 2865  
QY 2923 CTGTGCAACCTTTAGCGATTAAGCGTCTCTGTTTGCNACGGCGGGCTGGAACCGGAC 2982  
Db 2866 CTGTGCAACCTTTAGCGATTAAGCGTCTCTGTTTGCNACGGCGGGCTGGAACCGGAC 2925  
QY 2983 CTGAACGAGCGACTACACGCTAAGCGGCGCTTTACGGCGGACTGCACCAACCGCG 3042  
Db 2926 CTGAACGAGCGACTACACGCTAAGCGGCGCTTTACGGCGGCGACTGCACCAACCGCG 2985  
QY 3043 AAGACGGGGCACGCAATATGCCGACACCCGCCCTGGTTGCGCGCTGCGGCGCGGATGTC 3102  
Db 2986 AAGACGGGGCACGCAATATGCCGACACCCGCCCTGGTTGCGCGCGCTGCGGCGCGGATGTC 3045  
QY 3103 GAATTCGCAACGGCTGGAAACGGCTTGGCAGCTTACAGGTACGCGCGGTTCCAAACAGTAC 3162  
Db 3046 GAATTCGCAACGGCTGGAAACGGCTTGGCAGCTTACAGGTACGCGCGGTTCCAAACAGTAC 3105  
QY 3163 GGCACACAGCGGACGAGTGGCGGTAGGCTACCGGTTTC 3201  
Db 3106 GGCACACAGCGGACGAGTGGCGGTAGGCTACCGGTTTC 3144

RESULT 20  
LOCUS AX236407 4425 bp DNA linear PAT 26-SEP-2001  
DEFINITION Sequence 100 from Patent WO0164922.  
ACCESSION AX236407  
VERSION AX236407.1 GI:15796026  
KEYWORDS .  
SOURCE synthetic construct.  
ORGANISM synthetic construct.  
REFERENCE 1 (bases 1 to 4425)  
AUTHORS Arico,M.B., Comanducci,M.C., Galeotti,C.C., Masignani,I.V.C.,  
Guilianini,M.M. and Pizzo,M.C.  
TITLE Heterologous expression of neisserial proteins  
JOURNAL Patent: WO 0164922-A 100 07-SEP-2001;  
Chiron Spa (IT)

FEATURES  
source Location/Qualifiers  
1..4425  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="deltaG983-ORF46.1"

BASE COUNT 1165 a 1259 c 1219 g 782 t  
ORIGIN

Query Match 25.9%; Score 831; DB 6; Length 4425;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 1231; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 1963 CTGCTGGCTTCCTCGACAGCGTCTGAAACAAACAGCGGGCAGTGAAGGGCAGACACGTGTGCC 2022  
Db 1906 CTGCTGGCTTCCTCGACAGCGTCTGAAACAAACAGCGGGCAGTGAAGGGCAGACACGTGTGCC 1965  
QY 2023 TATTATGTCCTCGCGGCAATGCGGCACGGACTGCTTGGCAGCGGCACATTTCCCGGCC 2082  
Db 1966 TATTATGTCCTCGCGGCAATGCGGCACGGACTGCTTGGCAGCGGCACATTTCCCGGCC 2025  
QY 2083 GCCGCTCTGAAACACCGCGTAGAACAGGGCGGCAGCAATCTGAAAAACCTGATGTCGAA 2142  
Db 2026 GCCGCTCTGAAACACCGCGTAGAACAGGGCGGCAGCAATCTGAAAAACCTGATGTCGAA 2085  
QY 2143 CTGGATGCTTCGGAATCATTCGCAACACCCGAGAGCGGTTGAAACTGCGGCCCGGACCGC 2202  
Db 2086 CTGGATGCTTCGGAATCATTCGCAACACCCGAGAGCGGTTGAAACTGCGGCCCGGACCGC 2145  
QY 2203 ACAGATATGCCGGGCATCGGCCCTACGCGCAACTTTCCGCGCAGCGGCAGCGCTACAG 2262  
Db 2146 ACAGATATGCCGGGCATCGGCCCTACGCGCAACTTTCCGCGCAGCGGCAGCGCTACAG 2205  
QY 2263 CATGCGAATGCCCGCAGCGGTGTACGCATCTTTCAACAGTCTCGCGCTACCGCTCTATGCC 2322  
Db 2206 CATGCGAATGCCCGCAGCGGTGTACGCATCTTTCAACAGTCTCGCGCTACCGCTCTATGCC 2265  
QY 2323 GACATACCGCGCCCATGCGCATATGAGGACGCGCGGCTGAAAGCGGTATCGGACGGG 2382  
Db 2266 GACATACCGCGCCCATGCGCATATGAGGACGCGCGGCTGAAAGCGGTATCGGACGGG 2325  
QY 2383 TTGGACCAACAGCGTACGCGTCTGCGCTCATCGCGCAAAACCAACAGACCGCTGGAAG 2442  
Db 2326 TTGGACCAACAGCGCGCGGTCTGCGCTCATCGCGCAAAACCAACAGACCGGTGGAAG 2385  
QY 2443 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGCAAACTACCCAAACCTCGCGCATTTGCC 2502  
Db 2386 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGCAAACTACCCAAACCTCGCGCATTTGCC 2445  
QY 2503 GCGAAACCGGGGAAATACGACAGCAGCGCCACACTGGGGCATGGGACACAGACATGG 2562  
Db 2446 GCGAAACCGGGGAAATACGACAGCAGCGCCACACTGGGGCATGGGACAGCATGG 2505  
QY 2563 AGCGAAACAGTGCATAATGCAAAACCCGACAGCATTAGTCTCTTTGCGAGGCATACGGCAC 2622  
Db 2506 AGCGAAACAGTGCATAATGCAAAACCCGACAGCATTAGTCTCTTTGCGAGGCATACGGCAC 2565  
QY 2623 GATCGGGCGATATCGGCTATCTCAAGGCCCTGTTCTCCTACGGAGCGCTACAAAAACAGC 2682  
Db 2566 GATCGGGCGATATCGGCTATCTCAAGGCCCTGTTCTCCTACGGAGCGCTACAAAAACAGC 2625  
QY 2683 ATCAGCCGACAGCAGCGGTGCGGACGAACATGCGGAAGCGGCGGTCAACGGCAGCGTATG 2742  
Db 2626 ATCAGCCGACAGCAGCGGTGCGGACGAACATGCGGAAGCGGCGGTCAACGGCAGCGTATG 2685  
QY 2743 CAGCTGGCGCACTGGCGGTGTCAACGTTTCCGCGCAACGGGAGATTGACGGTTC 2802  
Db 2686 CAGCTGGCGCACTGGCGGTGTCAACGTTTCCGCGCAACGGGAGATTGACGGTTC 2745  
QY 2803 GAAGCGGTCTGCGCTACGACCTGCTCAAAACAGGATGCAATCCCGGCAAAAGGCAGTGT 2862  
Db 2746 GAAGCGGTCTGCGCTACGACCTGCTCAAAACAGGATGCAATCCCGGCAAAAGGCAGTGT 2805  
QY 2863 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGCAGCAGTGTGCGGACTCGCGGGTCTGAAG 2922  
Db 2806 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGCAGCAGTGTGCGGACTCGCGGGTCTGAAG 2865  
QY 2923 CTGTGCAACCTTTAGCGGATTAAGCGTCTCTGTTTGCNACGGCGGGCTGGAACCGGAC 2982  
Db 2866 CTGTGCAACCTTTAGCGGATTAAGCGTCTCTGTTTGCNACGGCGGGCTGGAACCGGAC 2925  
QY 2983 CTGAACGAGCGACTACACGTTAAGCGGCGCTTTACGGCGGCGACTGCACCAACCGCG 3042

Db 2926 CTGAACGGACGCGACTACAGGTAACGGGGGCTTTACGGCGGACTCGAGCAACCGCG 2985  
QY 3043 AAGACGGGGCAGCAATATATGCGCACACCGGCTGCTGCGGTCTGGCGGGATGTC 3102  
Db 2986 AAGACGGGGCAGCAATATATGCGCACACCGGCTGCTGCGGTCTGGCGGGATGTC 3045  
QY 3103 GAATTCTGGCAACGGCTGGAACGGCTTGGCACGTTTACAGCTACGCGGTTTCCAAACAGTAC 3162  
Db 3046 GAATTCTGGCAACGGCTGGAACGGCTTGGCACGTTTACAGCTACGCGGTTTCCAAACAGTAC 3105  
QY 3163 GGCNAACACAGCGGACAGTGGGGTAGGCTACCGGTTT 3201  
Db 3106 GGCNAACACAGCGGACAGTGGGGTAGGCTACCGGTTT 3144

RESULT 21  
AX239715  
LOCUS AX239715 4425 bp DNA linear PAT 26-SEP-2001  
DEFINITION Sequence 17 from Patent WO0164920.  
ACCESSION AX239715  
VERSION AX239715.1 GI:15797369  
KEYWORDS  
SOURCE synthetic construct.  
ORGANISM  
REFERENCE 1 (bases 1 to 4425)  
AUTHORS Arico, M.B., Comanducci, M., Galeotti, C., Massignani, V., Giuliani, M.M.  
TITLE Hybrid expression of neisserial proteins  
JOURNAL Patent: WO 0164920-A 17 07-SEP-2001;  
Chiron Spa (IT)  
FEATURES  
source  
1..4425  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="deltaG983-Orf46.1"

BASE COUNT 1165 a 1259 c 1219 g 782 t  
ORIGIN

Query Match 25.9%; Score 831; DB 6; Length 4425;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 1231; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1963 CTGCTGGCTTCCCTCGACAGCGTCGAAAAACAGCGGGCAGTGAAAGCGCACAGCTGTCC 2022  
Db 1906 CTGCTGGCTTCCCTCGACAGCGTCGAAAAACAGCGGGCAGTGAAAGCGCACAGCTGTCC 1965  
QY 2023 TATTATGTCGTCGCGCAATGCGGCACGAGCTGCTTTCGGCAGCGGCACATTCGCGCCCC 2082  
Db 1966 TATTATGTCGTCGCGCAATGCGGCACGAGCTGCTTTCGGCAGCGGCACATTCGCGCCCC 2025  
QY 2083 GCGGGCTGAAACACGCGGTAGAACAGGGGGGAGCAATCTGNAACCTGTATGGTGCAG 2142  
Db 2026 GCGGGCTGAAACACGCGGTAGAACAGGGGGGAGCAATCTGNAACCTGTATGGTGCAG 2085  
QY 2143 CTGATGCTTCCGAATCATCCGAACACCGAGAGCGGTGAAACTGCGCGCGGACCGC 2202  
Db 2086 CTGATGCTTCCGAATCATCCGAACACCGAGAGCGGTGAAACTGCGCGCGGACCGC 2145  
QY 2203 ACAGATATGCGGGCATCCGCCCTACGGCGCAACTTTCGGCGAGCGGCGAGCGGTACAG 2262  
Db 2146 ACAGATATGCGGGCATCCGCCCTACGGCGCAACTTTCGGCGAGCGGCGAGCGGTACAG 2205  
QY 2263 CATGCGAATGCGCGGACAGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2322  
Db 2206 CATGCGAATGCGCGGACAGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2265  
QY 2323 GACAGTACCGCGCCCATGCGGATATGAGGAGCGCGGTGAAACCGGTATCGGACGG 2382  
Db 2266 GACAGTACCGCGCCCATGCGGATATGAGGAGCGCGGTGAAACCGGTATCGGACGG 2325  
QY 2383 TTGACACACACGCTACGGGTCTGCGGTATCGCGCAACCCCAACAGGACGGTGAACG 2442

Db 2326 TTGACCAACAAGCGACGGGTCTCGCGTTCATCGCGCAACCCCAACAGGACGGTGAACG 2385  
QY 2443 TGGCAACAGGGCGGTGTGAAGGCAAAATGCGCGGAGTACCCCAACCGTGGCATTTGCC 2502  
Db 2386 TGGCAACAGGGCGGTGTGAAGGCAAAATGCGCGGAGTACCCCAACCGTGGCATTTGCC 2445  
QY 2503 GCGAAAAACCGCGCAAAATACGACAGCGCGCCACACTGGGCACTGGGACACAGCACATGG 2562  
Db 2446 GCGAAAAACCGCGCAAAATACGACAGCGCGCCACACTGGGCACTGGGACACAGCACATGG 2505  
QY 2563 AGCGAAAAACAGTGCAAATGCAAAACCGACAGCATTAAGTCTGTGTTGAGGCATACGCGAC 2622  
Db 2506 AGCGAAAAACAGTGCAAATGCAAAACCGACAGCATTAAGTCTGTGTTGAGGCATACGCGAC 2565  
QY 2623 GATCGGGCGGATATCGGCTATCTCAAAAGGCGCTGTCTCTACGACGCTACAAAAACAGC 2682  
Db 2566 GATCGGGCGGATATCGGCTATCTCAAAAGGCGCTGTCTCTACGACGCTACAAAAACAGC 2625  
QY 2683 ATCAGCGCGACGACCGGTGCGGAGCAACATGCGGAAGGCGAGCTCAACGGCACGCTGATG 2742  
Db 2626 ATCAGCGCGACGACCGGTGCGGAGCAACATGCGGAAGGCGAGCTCAACGGCACGCTGATG 2685  
QY 2743 CAGCTGGCGCACTGGCGGTGTCAACGTTCCGTTTCCGCAACCGGAGATTTGACGGTC 2802  
Db 2686 CAGCTGGCGCACTGGCGGTGTCAACGTTTCCGTTTCCGCAACCGGAGATTTGACGGTC 2745  
QY 2803 GAAGGCGGTCTGGCTACGACCTGCTCAACAGGATGATTCGCCGAAAAAGGAGGAGTGT 2862  
Db 2746 GAAGGCGGTCTGGCTACGACCTGCTCAACAGGATGATTCGCCGAAAAAGGAGGAGTGT 2805  
QY 2863 TTGGGCTGGAGCGGCAACAGCTCACTGAAGGACACTGTCGGGACTCGGGGTCTGAAG 2922  
Db 2806 TTGGGCTGGAGCGGCAACAGCTCACTGAAGGACACTGTCGGGACTCGGGGTCTGAAG 2865  
QY 2923 CTGTGCGCAACCTTGTAGCGATAAAGCGCTCCTGTTTCAACGGCGGCGGTGGAACGCGAC 2982  
Db 2866 CTGTGCGCAACCTTGTAGCGATAAAGCGCTCCTGTTTCAACGGCGGCGGTGGAACGCGAC 2925  
QY 2983 CTGAACGGACGCGACTACAGGTAACGGCGGCTTTACCGCGGCGACTCGAGCAACCGCG 3042  
Db 2926 CTGAACGGACGCGACTACAGGTAACGGCGGCTTTACCGCGGCGACTCGAGCAACCGCG 2985  
QY 3043 AAGACGGGGCAGCAATATATGCGCACACCGCGCTGCTGCGGTCTGGCGCGGATGTC 3102  
Db 2986 AAGACGGGGCAGCAATATATGCGCACACCGCGCTGCTGCGGTCTGGCGCGGATGTC 3045  
QY 3103 GAATTCTGGCAACGGCTGGAACGGCTTGGCACGTTTACAGCTACGCGGTTTCCAAACAGTAC 3162  
Db 3046 GAATTCTGGCAACGGCTGGAACGGCTTGGCACGTTTACAGCTACGCGGTTTCCAAACAGTAC 3105  
QY 3163 GGCNAACACAGCGGACAGTGGGGTAGGCTACCGGTTT 3201  
Db 3106 GGCNAACACAGCGGACAGTGGGGTAGGCTACCGGTTT 3144

RESULT 22  
NME277537  
LOCUS NME277537 3254 bp DNA linear BCT 24-JUL-2002  
DEFINITION Neisseria meningitidis ausp gene for autotransporter serine protease.  
ACCESSION AJ277537  
VERSION AJ277537.2 GI:13539237  
KEYWORDS ausp gene; autotransporter serine protease.  
SOURCE Neisseria meningitidis.  
ORGANISM Neisseria meningitidis  
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;  
Neisseria.  
REFERENCE 1  
Turner, D.P., Wooldridge, K.G. and Ala'Aldeen, D.A.  
TITLE Autotransported Serine Protease A of Neisseria meningitidis: an  
Immunogenic, Surface-Exposed Outer Membrane, and Secreted Protein  
JOURNAL Infect. Immun. 70 (8), 4447-4461 (2002)

MEDLINE 22112893  
REFERENCE 2  
AUTHORS Ala'Aldeen,D.A.A.  
TITLE Direct Submission  
JOURNAL Submitted (19-APR-2000) Ala'Aldeen D.A.A., Microbiology, University of Nottingham, Meningococcal Research Group, Division of Microbiology, University Hospital, Nottingham NG7 2UH, UNITED KINGDOM

REMARK Revised by [3]  
AUTHORS 3 (bases 1 to 3254)  
Ala'Aldeen,D.A.A.  
TITLE Direct Submission  
JOURNAL Submitted (02-APR-2001) Ala'Aldeen D.A.A., Microbiology, University of Nottingham, Meningococcal Research Group, Division of Microbiology, University Hospital, Nottingham NG7 2UH, UNITED KINGDOM

COMMENT On Apr 3, 2001 this sequence version replaced gi:7649689.  
FEATURES  
source Location/Qualifiers  
1..3254  
/organism="Neisseria meningitidis"  
/strain="B:15:P1.16"  
/isolate="SD"  
/db\_xref="Taxon:487"  
/country="United Kingdom:England"  
1..3254  
/gene="ausp"  
/pseudo  
1..3254  
/gene="ausp"  
/function="unknown"  
/pseudo  
/codon\_start=1  
/transl\_table=11  
/product="autotransporter serine protease"

BASE COUNT 845 a 922 c 917 g 570 t

ORIGIN  
Query Match 24.4%; Score 783; DB 1; Length 3254;  
Best Local Similarity 99.3%; Pred. No. 0;  
Matches 1233; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1963 CTGCTGGCTTCCTCGACAGCTCGAAGGAGGCGGCGGAGTGAAGGCGACACGCTGTGC 2022  
Db CTGCTGGCTTCCTCGACAGCTCGAAGGAGGCGGCGGAGTGAAGGCGACACGCTGTGC 2072

QY 2023 TATTATGTCGTCGGGCAATCGGCACGACTGCTTCGGCAGCGCAGATTCGCGGCC 2082  
Db TATTATGTCGTCGGGCAATCGGCACGACTGCTTCGGCAGCGCAGATTCGCGGCC 2132

QY 2083 GCGGTCTGAAACAGCCCTAGAACAGGCGGCGGAGCAATCTGGAAGCACTGATGTCGAA 2142  
Db GCGGTCTGAAACAGCCCTAGAACAGGCGGCGGAGCAATCTGGAAGCACTGATGTCGAA 2192

QY 2143 CTGGATGCTCCGAATATATCGGCAACACCCGAGACGGTTGAAACTCGCGCCGCCGACCCG 2202  
Db CTGGATGCTCCGAATATATCGGCAACACCCGAGACGGTTGAAACTCGCGCCGCCGACCCG 2252

QY 2203 ACAGATATCCGGGATTCGCCCTACGGCGCAACTTCCGGCGAGCGCAGCCGCTACAG 2262  
Db ACAGATATCCGGGATTCGCCCTACGGCGCAACTTCCGGCGAGCGCAGCCGCTACAG 2312

QY 2263 CATGCGAATGCGCGGACGGTGTAGCATCTTCAACAGTCTCCGCGTACCGCTCTATGCC 2322  
Db CATGCGAATGCGCGGACGGTGTAGCATCTTCAACAGTCTCCGCGTACCGCTCTATGCC 2372

QY 2323 GACAGTACCGCGCCCATCGGATATGACGAGGACCGCGGCTGAAAGCCGATATCGGACGGG 2382  
Db GACAGTACCGCGCCCATCGGATATGACGAGGACCGCGGCTGAAAGCCGATATCGGACGGG 2432

QY 2383 TTGGACCACACGCTACGGGTCTGGCGTCAATCGGCGAACCACAGACGCGGTGAAGC 2442  
Db TTGGACCACACGCGGCTGCGCGTCAATCGGCGAACCACAGACGCGGTGAAGC 2492

QY 2443 TGGGAACAGGCGGTGTTGAAGCAAAATGCGGCGAGTACCAAAACCGTCGGCATTTGCC 2502  
Db TGGGAACAGGCGGTGTTGAAGCAAAATGCGGCGAGTACCAAAACCGTCGGCATTTGCC 2552

QY 2503 GCGAAACCGGCGAAATACGACAGCAGCCGCCACACTGGGGATGGGACACACATGG 2562  
Db GCGAAACCGGCGAAATACGACAGCAGCCGCCACACTGGGGATGGGACACACATGG 2612

QY 2563 AGCGAAACAGTGCATAATGCAAAAACCGACACGATTAGTCTGTTTGCAGGCATACGGCAC 2622  
Db AGCGAAACAGTGCATAATGCAAAAACCGACACGATTAGTCTGTTTGCAGGCATACGGCAC 2672

QY 2623 GATGCGGCGGATATCGGCTATCTCAAAGGCGCTGTTCTCTCTACGAGCGGTACAAAACAGC 2682  
Db GATGCGGCGGATATCGGCTATCTCAAAGGCGCTGTTCTCTCTACGAGCGGTACAAAACAGC 2732

QY 2683 ATCAGCCGAGCACCGGTGCGGACGCAACATGCGGAAGGAGCGGTCAACGGCACGCTGATG 2742  
Db ATCAGCCGAGCACCGGTGCGGACGCAACATGCGGAAGGAGCGGTCAACGGCACGCTGATG 2792

QY 2743 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGGCGCAACGGAGATTGACGGTC 2802  
Db CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGGCGCAACGGAGATTGACGGTC 2852

QY 2803 GAAAGCGGTCTCGCTACGACCTGTCAAAACAGGATGATTCGCGGAAAGGCGAGTGTCT 2862  
Db GAAAGCGGTCTCGCTACGACCTGTCAAAACAGGATGATTCGCGGAAAGGCGAGTGTCT 2912

QY 2863 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGCGACACTGGTTCGGACTCGCGGGTCTGAAG 2922  
Db TTGGGCTGGAGCGGCAACAGCCTCACTGAAGCGACACTGGTTCGGACTCGCGGGTCTGAAG 2972

QY 2923 CTGTGCGCAACCTTTCAGCGATAAAGCGCTCTCTGTTTGAACGGCGGCGGTGAACGCGAC 2982  
Db CTGTGCGCAACCTTTCAGCGATAAAGCGCTCTCTGTTTGAACGGCGGCGGTGAACGCGAC 3032

QY 2983 CTGAACGCGGCGACTACACGGTAACGGGCGGCTTTACCGGCGGCGACTGCACAAACCGCG 3042  
Db CTGAACGCGGCGACTACACGGTAACGGGCGGCTTTACCGGCGGCGACTGCACAAACCGCG 3092

QY 3043 AAGACGGGCGCACGCAATATGCCGCACACCCCGCTGGTTGCGGGTCTGGGCGCGGATGTC 3102  
Db AAGACGGGCGCACGCAATATGCCGCACACCCCGCTGGTTGCGGGTCTGGGCGCGGATGTC 3152

QY 3103 GAATTCGCGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCGGTTCCAAACAGTAC 3162  
Db GAATTCGCGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCGGTTCCAAACAGTAC 3212

QY 3163 GGCACACGAGCGGACGAGTCGCGCTAGGCTACCGGTTCTGA 3204  
Db GGCACACGAGCGGACGAGTCGCGCTAGGCTACCGGTTCTGA 3254

RESULT 23  
AF169448/c 600 bp DNA linear BCT 09-AUG-2000  
LOCUS AF169448  
DEFINITION Neisseria meningitidis strain Z2491 clone Cm024 unknown sequence.  
ACCESSION AF169448  
VERSION AF169448.1 GI:9754656  
KEYWORDS  
SOURCE Neisseria meningitidis.  
ORGANISM Neisseria meningitidis  
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;  
Neisseria.  
REFERENCE 1 (bases 1 to 600)  
AUTHORS Perrin,A., Nassif,X. and Tinsley,C.R.  
TITLE Identification of regions of the chromosome of Neisseria meningitidis and Neisseria gonorrhoeae which are specific to pathogenic Neisseriae  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 600)  
AUTHORS Perrin,A., Nassif,X. and Tinsley,C.R.  
TITLE Direct Submission

JOURNAL Submitted (16-JUL-1999) Necker Medicine Faculty, INSERM U411, 156  
rue de Vaugirard, Paris 75015, France

## FEATURES

Source  
1..600  
Location/Qualifiers  
/organism="Neisseria meningitidis"  
/strain="Z2491"  
/db\_xref="taxon:487"  
/clone="Cm024"

BASE COUNT 111 a 159 c 144 g 184 t 2 others

## ORIGIN

Query Match 10.0%; Score 321; DB 1; Length 600;  
Best Local Similarity 100.0%; Pred. No. 6.1e-169;  
Matches 321; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 565 GAGCGCGTTATACAGACTGAAGCAACCGCACGGATATCGCCACCTAAAGAAATCGGA 624  
DB 341 GAGCGCGTTATACAGACTGAAGCAACCGCACGGATATCGCCACCTAAAGAAATCGGA 282  
QY 625 CACATCGATGTGCTCTCCCATATATTATGGCGGCGTTCCGTGGACGCGACACCTGCAGGC 684  
DB 281 CACATCGATGTGCTCTCCCATATATTATGGCGGCGTTCCGTGGACGCGACACCTGCAGGC 222  
QY 685 GGTATTCGCGCCGATCGGACGCTACACATATGAATACGCATGATGGAACCAAGACGAA 744  
DB 221 GGTATTCGCGCCGATCGGACGCTACACATATGAATACGCATGATGGAACCAAGACGAA 162  
QY 745 ATAATCTCTGCACCATCCGCAATGATGGTCAAGCTGGCGCAAGCTGCGTGCATC 804  
DB 161 ATAATCTCTGCACCATCCGCAATGATGGTCAAGCTGGCGCAAGCTGCGTGCATC 102  
QY 805 GTCAATAACAGTTTTTGGAAACAACATCGAGGCGAGGCACTGCCACCATTTCCAAATAGCC 864  
DB 101 GTCAATAACAGTTTTTGGAAACAACATCGAGGCGAGGCACTGCCACCATTTCCAAATAGCC 42  
QY 865 AATTCGGAGGAGCAGTACCGC 885  
DB 41 AATTCGGAGGAGCAGTACCGC 21

## RESULT 24

AF169473 286 bp DNA linear BCT 09-AUG-2000  
DEFINITION Neisseria meningitidis strain Z2491 clone Em085 unknown sequence.

## LOCUS

AF169473

VERSION AF169473.1 GI:9754681

KEYWORDS Neisseria meningitidis.

SOURCE Neisseria meningitidis

ORGANISM Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;

Neisseria.

REFERENCE 1 (bases 1 to 286)

Perrin,A., Nassif,X. and Tinsley,C.R.

Identification of regions of the chromosome of Neisseria

meningitidis and Neisseria gonorrhoeae which are specific to

pathogenic Neisseriae

Unpublished

REFERENCE 2 (bases 1 to 286)

Perrin,A., Nassif,X. and Tinsley,C.R.

Direct Submission

Submitted (16-JUL-1999) Necker Medicine Faculty, INSERM U411, 156

rue de Vaugirard, Paris 75015, France

Location/Qualifiers

1..286

/organism="Neisseria meningitidis"

/strain="Z2491"

/db\_xref="taxon:487"

/clone="Em085"

BASE COUNT 88 a 73 c 63 g 62 t

ORIGIN

## Query Match

Best Local Similarity 5.5%; Score 177; DB 1; Length 286;

Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 865 AATTCGGAGGAGCAGTACCGCCAAAGCGTTGCTCGCTATTTCGCGGTGATAAAGACAGAC 924  
DB 1 AATTCGGAGGAGCAGTACCGCCAAAGCGTTGCTCGCTATTTCGCGGTGATAAAGACAGAC 60

QY 925 GAGGGTATCGCCTGATGCAACAGAGCGATTACGGCAACTTGTCTTACCACATCCGTAAT 984  
DB 61 GAGGGTATCGCCTGATGCAACAGAGCGATTACGGCAACTTGTCTTACCACATCCGTAAT 120

QY 985 AAAAAACATGCTTTTTCATTTTTCGGCAAGCAATGAGCACAAGCTCAGCCCAACACA 1041  
DB 121 AAAAAACATGCTTTTTCATTTTTCGGCAAGCAATGAGCACAAGCTCAGCCCAACACA 177

## RESULT 25

A68924

LOCUS

DEFINITION

Accession

Version

Keywords

Source

Organism

Reference

Authors

Title

Journal

Comment

Features

Source

Base Count

Origin

Query Match

Best Local Similarity

Matches

QY 865 AATTCGGAGGAGCAGTACCGCCAAAGCGTTGCTCGCTATTTCGCGGTGATAAAGACAGAC 924

DB 1 AATTCGGAGGAGCAGTACCGCCAAAGCGTTGCTCGCTATTTCGCGGTGATAAAGACAGAC 60

QY 925 GAGGGTATCGCCTGATGCAACAGAGCGATTACGGCAACTTGTCTTACCACATCCGTAAT 984

DB 61 GAGGGTATCGCCTGATGCAACAGAGCGATTACGGCAACTTGTCTTACCACATCCGTAAT 120

QY 985 AAAAAACATGCTTTTTCATTTTTCGGCAAGCAATGAGCACAAGCTCAGCCCAACACA 1041

DB 121 AAAAAACATGCTTTTTCATTTTTCGGCAAGCAATGAGCACAAGCTCAGCCCAACACA 177

## RESULT 26

AX236815

LOCUS

DEFINITION

Accession

Version

Keywords

Source

Organism

Reference

Authors

Title

Journal

Query Match

Best Local Similarity

Matches

QY 865 AATTCGGAGGAGCAGTACCGCCAAAGCGTTGCTCGCTATTTCGCGGTGATAAAGACAGAC 924

DB 1 AATTCGGAGGAGCAGTACCGCCAAAGCGTTGCTCGCTATTTCGCGGTGATAAAGACAGAC 60

QY 925 GAGGGTATCGCCTGATGCAACAGAGCGATTACGGCAACTTGTCTTACCACATCCGTAAT 984

DB 61 GAGGGTATCGCCTGATGCAACAGAGCGATTACGGCAACTTGTCTTACCACATCCGTAAT 120

QY 985 AAAAAACATGCTTTTTCATTTTTCGGCAAGCAATGAGCACAAGCTCAGCCCAACACA 1041

DB 121 AAAAAACATGCTTTTTCATTTTTCGGCAAGCAATGAGCACAAGCTCAGCCCAACACA 177

## RESULT 26

AX236815

LOCUS

DEFINITION

Accession

Version

Keywords

Source

Organism

Reference

Authors

Title

Journal

Query Match

Best Local Similarity

Matches

QY 865 AATTCGGAGGAGCAGTACCGCCAAAGCGTTGCTCGCTATTTCGCGGTGATAAAGACAGAC 924

DB 1 AATTCGGAGGAGCAGTACCGCCAAAGCGTTGCTCGCTATTTCGCGGTGATAAAGACAGAC 60

QY 925 GAGGGTATCGCCTGATGCAACAGAGCGATTACGGCAACTTGTCTTACCACATCCGTAAT 984

DB 61 GAGGGTATCGCCTGATGCAACAGAGCGATTACGGCAACTTGTCTTACCACATCCGTAAT 120

QY 985 AAAAAACATGCTTTTTCATTTTTCGGCAAGCAATGAGCACAAGCTCAGCCCAACACA 1041

DB 121 AAAAAACATGCTTTTTCATTTTTCGGCAAGCAATGAGCACAAGCTCAGCCCAACACA 177

```

FEATURES
  source
    Chiron Spa (IT)
    Location/Qualifiers
      1..50
      /organism="synthetic construct"
      /db_xref="taxon:32630"
      /note="Oligonucleotide"
BASE COUNT
  15 a 20 c 7 g 8 t
ORIGIN
  Query Match
    1..18: Score 35; DB 6; Length 50;
  Best Local Similarity
    100.0%; Pred. No. 3.1e-07;
  Matches
    35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 CGACGACCCCAACCTTCCTACAAAAAATTCAA 38
|||||
Db 16 CGACGACCCCAACCTTCCTACAAAAAATTCAA 50

RESULT 27
NGOPAK
LOCUS
  NGOPAK 1434 bp DNA linear BCT 10-APR-1995
DEFINITION
  N.gonorrhoeae opak gene for opacity protein.
ACCESSION
  X52364 S36068 S75784
VERSION
  X52364.1 GI:48702
KEYWORDS
  opacity protein; opak gene.
SOURCE
  Neisseria gonorrhoeae.
ORGANISM
  Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
  Neisseria.
REFERENCE
  1 (bases 1 to 1434)
  Bhat,K.
  Direct Submission
  Submitted (31-MAR-1990) Bhat K., National Institutes of Health,
  Rocky Mountain Laboratory, Hamilton 59840, U S A
  2 (bases 1 to 1318)
  Bhat,K.S., Gibbs,C.P., Barrera,O., Morrison,S.G., Jahnig,F.,
  Stern,A., Kupsch,E.M., Meyer,T.F. and Swanson,J.
  The opacity proteins of Neisseria gonorrhoeae strain MS11 are
  encoded by a family of 11 complete genes
  Mol. Microbiol. 5 (8), 1889-1901 (1991)
  92114767
  1815562
  3 (bases 1 to 1434)
  Bhat,K.S., Gibbs,C.P., Barrera,O., Morrison,S.G., Jahnig,F.,
  Stern,A., Kupsch,E.M., Meyer,T.F. and Swanson,J.
  The opacity proteins of Neisseria gonorrhoeae strain MS11 are
  encoded by a family of 11 complete genes
  Mol. Microbiol. 6 (8), 1073-1076 (1992)
  92261323
  1584024
  The cloned gene is in its non-expressed state. The change in number
  of the coding repeat unit has caused the atg start codon in the
  signal peptide to be out of frame with the rest of its coding
  sequence.
FEATURES
  source
    Location/Qualifiers
      1..1434
      /organism="Neisseria gonorrhoeae"
      /strain="MS11"
      /db_xref="taxon:485"
      218..243
      repeat_unit
      repeat_unit
      -35_signal
      -10_signal
      RBS
      gene
      /gene="opak"
      /note="The cloned gene is in its non-expressed state. The
      change in the number of coding repeat unit has caused the
      atg start codon to be out of frame with the rest of its
      coding sequence"
      /pseudo

FEATURES
  source
    Location/Qualifiers
      1..37497
      /organism="Drosophila melanogaster"
      /db_xref="taxon:7227"
BASE COUNT
  10676 a 8030 c 7947 g 10844 t
ORIGIN
  Query Match
    0.8%; Score 25; DB 2; Length 37497;
  Best Local Similarity
    100.0%; Pred. No. 0.15;
  Matches
    25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 143 GCAGCAACAGCAGACCAACACAGC 167
|||||
Db 26162 GCAGCAACAGCAGACCAACACAGC 26138

RESULT 29
AC009749/c
LOCUS
  AC009749 161278 bp DNA linear INV 28-JUL-2001
DEFINITION
  Drosophila melanogaster, chromosome 2L, region 25B-25C, BAC clone
  BACR40104, complete sequence.
ACCESSION
  AC009749
VERSION
  AC009749.8 GI:15027726
KEYWORDS
  HTG.
SOURCE
  Drosophila melanogaster.
ORGANISM
  Drosophila melanogaster.
  /codon_start=1
  /transl_table=11
  /product="opacity protein"
  /db_xref="PID:el45998"
  490..524
  /note="signal peptide coding sequence"
  /rpt_type=TANDEM
  539..1252
  /gene="opak"
  /product="opacity protein"
  1269..1294
  343 g 283 t
BASE COUNT
  383 a 425 c
ORIGIN
  Query Match
    1.0%; Score 33; DB 1; Length 1434;
  Best Local Similarity
    100.0%; Pred. No. 4.5e-06;
  Matches
    33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3146 CCGGTTCCAAACAGTACGGCAACACACACGGGAC 3178
|||||
Db 30 CCGGTTCCAAACAGTACGGCAACACACACGGGAC 62

RESULT 28
AC017356/c
LOCUS
  AC017356 37497 bp DNA linear HTG 09-DEC-1999
DEFINITION
  Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***, in Ordered
  pieces.
ACCESSION
  AC017356
VERSION
  AC017356.1 GI:6553630
KEYWORDS
  HTG; HTGS-PHASE2.
SOURCE
  Drosophila melanogaster.
ORGANISM
  Drosophila melanogaster.
  Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
  Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
  Ephydroidea; Drosophilidae; Drosophila.
  1 (bases 1 to 37497)
  Adams,M. and Venter,J.C.
  Direct Submission
  Submitted (09-DEC-1999) Celera Genomics, 45 West Gude Drive,
  Rockville, MD, USA
  This sequence was identified as CDM:10210411 by the submitter.
  For more information on this record e-mail to fly@celera.com.
  * NOTE: This is a 'working draft' sequence.
  * This sequence will be replaced
  * by the finished sequence as soon as it is available and
  * the accession number will be preserved.
FEATURES
  source
    Location/Qualifiers
      1..37497
      /organism="Drosophila melanogaster"
      /db_xref="taxon:7227"
BASE COUNT
  10676 a 8030 c 7947 g 10844 t
ORIGIN
  Query Match
    0.8%; Score 25; DB 2; Length 37497;
  Best Local Similarity
    100.0%; Pred. No. 0.15;
  Matches
    25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 143 GCAGCAACAGCAGACCAACACAGC 167
|||||
Db 26162 GCAGCAACAGCAGACCAACACAGC 26138

RESULT 29
AC009749/c
LOCUS
  AC009749 161278 bp DNA linear INV 28-JUL-2001
DEFINITION
  Drosophila melanogaster, chromosome 2L, region 25B-25C, BAC clone
  BACR40104, complete sequence.
ACCESSION
  AC009749
VERSION
  AC009749.8 GI:15027726
KEYWORDS
  HTG.
SOURCE
  Drosophila melanogaster.
ORGANISM
  Drosophila melanogaster.

```



Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 161278)

Celniker,S.E., Adams,M.D., Krommiller,B., Tyler,D., Wan,K.H., Holt,R.A., Evans,C.A., Gocayne,J.D., Ananatiades,P.G., Rogers,Y., An,H., Baldwin,D., Banzon,J., Beeson,K.Y., Busan,D.A., Carlson,J.W., Center,A., Champe,M., Davenport,L.B., Dietz,S.M., Dodson,K., Dorsett,V., Doup,L.E., Doyle,C., Dresnek,D., Farfan,D., Ferriera,S., Frise,E., Galle,R.F., Garg,N.S., George,R.A., Gonzalez,M., Houck,J., Hoskins,R.A., Hostin,D., Howland,T.J., Ibegwam,C., Jalali,M., Kruse,D., Li,P., Mattel,B., Moshrefi,A., McIntosh,T.C., Moy,M., Murphy,B., Nelson,K., Nelson,K.A., Nunoo,J., Pacleb,J., Paragas,V., Park,S., Patel,S., Pfeiffer,B., Phouanavong,S., Pittman,G.S., Puri,V., Richards,S., Scheeler,F., Stapleton,M., Strong,R., Svirskas,R., Tector,C., Williams,S.M., Zaveri,J.S., Smith,H.O., Rubin,G.M. and Venter,C.B.

Sequencing of Drosophila chromosome 2L, region 25B-25C

Unpublished

2 (bases 1 to 161278)

Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G., Butenhoff,C., Chavez,C., Chew,M., Ciesiolka,L., Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L., Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L., Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P., Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S., Pfeiffer,B., Poon,L., Sequeira,A., Sethi,H., Snir,E., Svirskas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and Rubin,G.M.

Direct Submission

Submitted (31-Aug-1999) Drosophila Genome Center, Lawrence Berkeley Laboratory, MS 64-121, Berkeley, CA 94720, USA

On Jul 28, 2001 this sequence version replaced gi:6980162.

Sequence submitted by:

Berkeley Drosophila Genome Project

Lawrence Berkeley National Laboratory, MS 64-121 Berkeley, CA 94720

This sequence was assembled using end sequences from a whole genome shotgun and from subclones of this BAC and its neighboring clones. For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive Web site (<http://www.fruitfly.org/sequence/>) or send email to [bdgpe@fruitfly.berkeley.edu](mailto:bdgpe@fruitfly.berkeley.edu).

Location/Qualifiers

1. 161278

/organism="Drosophila melanogaster"

/strain="v: cn bw sp"

/db\_xref="taxon:7227"

/chromosome="2L"

/map="25B-25C"

/clone="BACR40104 (Dl088)"

/clone.lib="RpCi-98 (Roswell Park Cancer Institute Drosophila melanogaster BAC library, partial EcoRI in pBAC3.6)"

BASE COUNT 44503 a 35245 c 35792 g 45738 t

ORIGIN

Query Match 0.8% Score 25; DB 3; Length 161278;

Best Local Similarity 100.0%; Pred. No. 0.15;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 143 GCACGACAGCAGACACACACG 167

|||||

Db 119005 GCACGACAGCAGACACACACG 118981

|||||

RESULT 30

AE003575

LOCUS

AE003575 327209 bp DNA linear INV 04-OCT-2000

DEFINITION

Drosophila melanogaster genomic scaffold 142000013386046 section 16 of 16, complete sequence.

ACCESSION

AE003575 AE002638

VERSION

AE003575.2 GI:10727324

## KEYWORDS

SOURCE

ORGANISM

HTG.

Drosophila melanogaster.

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 327209)

## REFERENCE

AUTHORS

Adams,M.D., Celniker,S.E., Holt,R.A., Evans,C.A., Gocayne,J.D., Ananatiades,P.G., Scherer,S.E., Li,P.W., Hoskins,R.A., Galle,R.F., George,R.A., Lewis,S.E., Richards,S., Ashburner,M., Henderson,S.N., Sutton,G.G., Wortman,J.R., Yandell,M.D., Zhang,Q., Chen,L.X., Brandon,R.C., Rogers,Y.H., Blazej,R.G., Champe,M., Pfeiffer,B.D., Wan,K.H., Doyle,C., Baxter,E.G., Helt,G., Nelson,C.R., Gabor Miklos,G.L., Abril,J.F., Agbayani,A., An,H.J., Andrews-Pfannkoch,C., Baldwin,D., Ballew,R.M., Basu,A., Baxendale,J., Bayraktaroglu,L., Beasley,E.M., Beeson,K.Y., Benos,P.V., Bereman,B.P., Bhandari,D., Bolshakov,S., Borkova,D., Botchan,M.R., Bouck,J., Brokstein,P., Brottier,P., Burtis,K.C., Busam,D.A., Butler,H., Cadieu,E., Center,A., Chandra,I., Cherry,J.M., Cawley,S., Dahlke,C., Davenport,L.B., Davies,P., de Pablos,B., Delcher,A., Deng,Z., Mays,A.D., Dew,I., Dietz,S.M., Dodson,K., Doup,L.E., Downes,M., Dugan-Rocha,S., Dunkov,B.C., Dunn,P., Durbin,K.J., Evangelista,C.C., Ferraz,C., Ferrieres,S., Fleischmann,W., Fosler,C., Gabrilian,A.E., Garg,N.S., Gelbart,W.M., Glasser,K., Glodek,A., Gong,F., Gorrell,J.H., Gu,Z., Guan,P., Harris,M., Harris,N.L., Harvey,D., Heiman,T.J., Hernandez,J.R., Houck,J., Hostin,D., Houston,K.A., Howland,T.J., Wei,M.H., Ibegwam,C., Jalali,M., Kalush,F., Karpen,G.H., Ke,Z., Kennison,J.A., Ketchum,K.A., Kimmel,B.E., Kodira,C.D., Kraft,C., Kravitz,S., Kulp,D., Lai,Z., Lasko,P., Lei,Y., Levitsky,A.A., Li,J., Li,Z., Liang,Y., Lin,X., Liu,X., Liu,X., Mattel,B., McIntosh,T.C., McLeod,M.P., McPherson,D., Merkulov,G., Milshina,N.V., Mobarry,C., Morris,J., Moshrefi,A., Mount,S.M., Moy,M., Murphy,B., Murphy,L., Muzny,D.M., Nelson,D.L., Nelson,D.R., Nelson,K.A., Nixon,K., Nusskern,D.R., Pacleb,J.M., Palazzolo,M., Pittman,G.S., Pan,S., Pollard,J., Puri,V., Reese,M.G., Reinert,K., Remington,K., Saunders,R.D., Scheeler,F., Shen,H., Shue,B.C., Siden-Kiamos,I., Simpson,M., Skupski,M.P., Smith,T., Spier,E., Spradling,A.C., Stapleton,M., Strong,R., Sun,E., Svirskas,R., Tector,C., Turner,R., Venter,E., Wang,A.H., Wang,X., Wang,Z.Y., Wassarman,D.A., Weinstock,G.M., Weissbach,J., Williams,S.M., Woodage,T., Worley,K.C., Wu,D., Yang,S., Yao,Q.A., Ye,J., Yeh,R.F., Zaveri,J.S., Zhan,M., Zhang,G., Zhao,Q., Zheng,L., Zheng,X.H., Zhong,F.N., Zhong,W., Zhou,X., Zhu,S., Zhu,X., Smith,H.O., Gibbs,R.A., Myers,E.W., Rubin,G.M. and Venter,J.C.

The genome sequence of Drosophila melanogaster

Science 287 (5461), 2185-2195 (2000)

20196006

10731132

REFERENCE

2 (bases 1 to 327209)

Adams,M.D., Celniker,S.E., Gibbs,R.A., Rubin,G.M. and Venter,C.J.

Direct Submission

Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive, Rockville, MD, USA

On Oct 9, 2000 this sequence version replaced gi:7295650.

Location/Qualifiers

1. 327209

/organism="Drosophila melanogaster"

/db\_xref="taxon:7227"

/chromosome="2L"

<29815..>30955

/gene="CG11933"

/evidence=not\_experimental

/map="25A3-25A3"

/db\_xref="FLYBASE:FBan0011933"

/db\_xref="FLYBASE:FBan0031615"

join(<29815..29907,30773..>30955)

/gene="CG11933"

/product="CT35796"

/evidence=not\_experimental

/db\_xref="FLYBASE:FBan0011933"

/db\_xref="FLYBASE:FBan0031615"

join(29815..29907,30773..30955)

CDS



FEATURES	PASTEUR MERIEUX SERUMS VACC (FR)										
	source	1. .30									
		/organism="synthetic construct"									
		/db_xref="taxon:32630"									
BASE COUNT	ORIGIN	7 a	11 c	9 g	3 t						
							gene	CDS			
Query Match 0.7%; Score 23; DB 6; Length 30;											
Best Local Similarity 100.0%; Pred. No. 1.7;											
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;											
QY	3181	GTCGGCGTAGGCTACCGTTCTG 3203									
LOCUS	DEFINITION										
Db	30	GTCGCGTAGGCTACCGTTCTG 8									
RESULT 32											
AB014075/c											
LOCUS											
AB014075	14043 bp	DNA	linear	BCT 08-MAY-1999							
DEFINITION	Clostridium histolyticum genes for hypoxanthine-guanine phosphoribosyl-transferase (HGPRTase), GTPase and 12 ORFs, complete and partial cds.										
AB014075											
ACCESSION	AB014075.1	GI:3868863									
VERSION	rvuB; colH; Orf2u; Orf3u; Orf4u; Orf5u; Orf6u; hf1X; GTPase; hprT;										
KEYWORDS	hypoxanthine-guanine phosphoribosyl-transferase (HGPRTase); Orf9u.										
SOURCE	Clostridium histolyticum (strain:JCM 1403) vegetative cell DNA.										
ORGANISM	Clostridium histolyticum										
REFERENCE	Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;										
AUTHORS	1 (sites)										
	Matsushita,O., Jung,C.M., Katayama,S., Minami,J., Takahashi,Y. and Okabe,A.										
TITLE	Gene duplication and multiplicity of collagenases in Clostridium histolyticum										
JOURNAL	J. Bacteriol.	181 (3),	923-933	(1999)							
MEDLINE	99121032										
REFERENCE	2 (bases 1 to 14043)										
AUTHORS	Matsushita,O.										
TITLE	Direct Submission										
JOURNAL	Submitted (18-May-1998) Osamu Matsushita, Kagawa Medical University, Department of Microbiology; 1750-1 Ikenobe, Miki-cho, Kagawa 761-0793, Japan (E-mail:osamu@kms.ac.jp, Tel.:+81(87)891-2129, Fax:+81(87)891-2129)										
FEATURES	source	Location/Qualifiers									
		1. .14043									
		/organism="Clostridium histolyticum"									
		/strain="JCM 1403"									
CDS	CDS	/db_xref="taxon:1498"									
		/dev_stage="vegetative cell"									
		153. .800									
		/note="Orf9u"									
gene	CDS	/codon_start=1									
		/transl_table=11									
		/protein_id="BAA34252.1"									
		/db_xref="GI:3868864"									
translation="MKFNKAKLISLIICIIIIITFYVLFDSIVPTVAVSEAMR											
	AKTEIINGCIIIDEYKGFNNDIIKDDDEGNISLRIDLKLNKIASISLKGQK										
	KLNELRGIKLSLGYITKNILSYIGPDIAVKKEPTSNVETYSIFKSAGINQTSK										
	KIFIKFKIRIIIPKSDVDVEVYSEMPISETIIVGKIPDSAIQLDLNKAQYKNN"										
complement(818. .1345)											
	/gene="hprT"										
	Complement(818. .1345)										
	/gene="hprT"										
codon_start=1											
	/transl_table=11										
	/product="hypoxanthine-guanine phosphoribosyl-transferase (HGPRTase)"										
	/protein_id="BAA34253.1"										
db_xref="GI:3868865"											
	/translation="MEDKKRNILIGKEEIASKIQEIGAIISKDYEGKNLYVLSLRGS										

FLFTADLVRELQVTKIGFMTTSSYGNSENSSGTVKVVNDVPDNIIEGFDVLIVDDIMD  
TGITMEFVVVKKLKGANSVNTCVLLDKPERRVAKIKPDYTCFTEIPDLFVVGGLNYG  
DYRNIPVYFNWENN"  
1469. .3259  
/gene="hf1X"  
1469. .3259  
/gene="hf1X"  
/codon\_start=1  
/transl\_table=11  
/product="GTPase"  
/protein\_id="BAA34254.1"  
/db\_xref="GI:3868866"  
/translation="MIYGNIEGLKDSILERLEKIYKMPVWKDSIFTEEMVSVIEEVTL  
KINREISIAISRKGVSVAVGDSSTVDLPLIDIKENRLSGVRIITHPHNGPKLSAI  
DISALTKLKDCIIVAGYEENRRKVNVCYSIENDTCCDEIGPLTLEETISLRLE  
RVFVEKLLKQSEIYEBQGERAVLVGISESELEBELAEACNVETCSIFQRRERA  
DNAPLVSGKVDDEISLAQAYKGNVIFDEELTGAQVRNLEENLTGVKVIDRTLLLEI  
FASARSREAKLOVELAQRYRSARLIGLTLSGRTGGIGTRGCKKLELTKRRIR  
ERITELSRLEKKITSIREIQORDKRVGNIPKISLVGTYNTGKSLRNLKCELTVPKEN  
INKEVFANMLFATIDTTTRAIAPDGRACLTSDTVGFIKRLHDLVESPKSTLEE  
VNSKLLHVVDSSSEKLYEIDAVEMVLSQGVDPKECILVNLKIDKLTKELELDEL  
KLNKYKLVYISAKTGENLDLILNAICENLPQTLREITFLIPYTEOGKASFHRNAKV  
DENYEDDGTAKALVDDEVYNKCMETEVKC"  
3272. .4168  
/note="Orf6u"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="BAA34255.1"  
/db\_xref="GI:3868867"  
/translation="MPKTISEYQAEANEALENLTKDSVAAIVVFGSIVNGDIWEKSD  
IDMVISEKQEKIVNIYSDYDGIPIHVRLIHKDELMNLCSDGNKGSHIHRFSSSK  
VYSKTEIDRLYNEERFYSDIHREKNVTFLSKLLKVSSECKYSLMDSLVNAYTLCN  
IVLRDYSKLVNSSGCIINRDIYVMAISTDELSREVLQFNDXNLKCKICGRFRFI  
EKGKINRITIASIIQYLDRESNALSSEIRKDLGFKHYDINIIEILDSIYELNIIN  
RCKRGYKTLNEETLIDENYYIL"  
4180. .5580  
/note="Orf5u"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="BAA34256.1"  
/db\_xref="GI:3868868"  
/translation="MEIYIEFKDPKRYVQISQIKVIERIIIGIINDGEKLPISRKLS  
ELLNVKDTVINAYRKLLEEGYAIQIGSVYAKRDTGRFKKDYQNTIKNLSYKNO  
EKYIDLVEATISSYFPFKNFTVLNEVLDRDGSKAFYKEVLGEGLRKSNINNFWD  
GKRKSEELIVSGAOGGIDVLSKALINVRDNLIVKPYNGALSVEKPRRANIIEVPI  
ENDGIDLQDFEKILQKNIKKCPYANSFQNPFGVSYMKKKRIELELSDYDFYIIE  
DYLSELIYESITYNSFKSLDINDRVIYIKFSKIFLPGLRIGYLDICPKKFLDSTINC  
KINTDITTSMLQRALDLYINRGYKHEHIDLRNNYKETYNYVLTVEKLAKELSEYSFN  
LPKGLNLYLKLHNTNSIDFFYKCKEQNIITTPGVFFYKYNKEGLDYVKLGFSNLRTD  
OILSVTNILKEIFKTK"  
5599. .6273  
/note="Orf4u"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="BAA34257.1"  
/db\_xref="GI:3868869"  
/translation="MKKGLINFAPWDIDKFSSELSYVLFEGKSIDVLSKIRSRE  
EQSHIINGKIKFGLSRASNEKELEFELCTVAKEDRRKSGFLSVNLKAEELNTIY  
NYSMDWSKKERAIWIIGEAKAENAIISLLKASVHKNNLRRLAISALGKLENKIGES  
IMRALEDNDPQVLQYAIKSLQKINSTRGIEKIKKIYENTOKYKLDCTETVLPQVED  
6286. .6936  
/note="Orf3u"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="BAA34258.1"  
/db\_xref="GI:3868870"  
/translation="MSYFTIKSKANSEFQEKKSIFIGHVKRVNSEEAKQFTEETKEK  
YKDAHNVYAYIIGENMGIQRYSDGEPKGTGGIPILEVIKKNITDITVVVYTRFEGG  
ILLGASGUTRAYSKAAVNAIKEAGVVEKIIIGAKVNDIVDEYELLGKIQTYSOKNIYIE  
DIQYSDKVTITVYFLELDIADVEREVMNITSGNGILLRDDAKYFFKLEDRLYEEQ"  
7024. .7764  
/note="Orf2u"

```
/codon_start=1
/transl_table=11
/protein_id="BAA34259.1"
/db_xref="GI:3892649"
/transl_table="MSHSHKWHNQIAKSKVDKARGKFTKIGKEIAIAAKNGPNPD
ANKLRDVIKAKANHPKDSIERAIKKAAGELAGVNYBELLYEGYDPGLVAVQAL
TDNKRAGNVRHAFSKHGNGLSTGCVSMFQKGQIVIEKNDELDEEELMMALSC
GAEDFQSEDEYIITTSPEDFGVSRETLEKGELEAEVKSIPDTVTYVIDENTAGKF
QKMLDVLDEDEDDVDQVYHNAEFPPEGWEE"
complement(7932..8411)
/note="Orflu"
/codon_start=1
/transl_table=11
/protein_id="BAA34541.1"
/db_xref="GI:3892645"
/transl_table="MNHKKEKEYKLDTEKLENTELQYNVTORNATEPFLNKYDK
HFDGIYDVLSGPEPLSIDKNSGCGMPAFSPKRYIKERADSHGMSRIEVS
KNADSHGHVFNDDPIENGMRVCINSASLKFIAKDLKEEGYEFLPLFEKQDEL"
8715..11780
/gene="colH"
/contig="11780"
/codon_start=1
/transl_table=11
/protein_id="BAA34542.1"
/db_xref="GI:3892646"
/transl_table="MKRKLCSKRLMLATMTATFTVNSTPLPIYAAVDKNATAAVONE
SKRYTYSYKLTNYDVLVLLVKTIEIENLPDLFOYSSDAKEFYGNKTRMSFIMDEIGR
RAPYOTEDIHGIPITLVVVVAGPYLGFHNLKELNEIKRSFKERVIPSILAIQKNPF
KLGEVDQKIVSATGLLAGNETAPPEVVNFTPILODCIRNIDRYALDDLUKSKALNF
LAAPYDITEYLRATKEKPEPTWYKIDGFIENLKLALYKINDNSWIIDNGIYH
IAPLGKLNHNKIGIETITEVMKYPIYLSMOHQLQSAQIKRHVDSKDAEGNKIPIDRF
KKEGKCYKCTMYTFDQGVILIKAGARVEBEKVRKYIPEQTEFTFYERAEQESTYLE
EGNPDILITWYNSPEYKLSVLYGYDTNNGMWIPEQTEFTFYERAEQESTYLE
ELPHEVTHYLGRAVPGOMGRKLYDNDRLTWYBEGAELEFAGSTRTSGILPRKSI
VSNHNTNRNRYKLSDTVSKYSGSEFFVNYACMFMDYNYKDMGLINKLNDLAKNN
VDGDNTRIDLSNLYNDKIYQHMQRIDNYENLVFVADDDLYVRHAYKNPNEYI
SESEVAKLDKADSEVKSSQYFSTFTLRGSTGGASKLEDQKAMNKFIDDSLKLID
TYSNGKLTATYNTKVDSSNRVTYDVVHGYLPNEDGSKNSLPGKINGTYKGTG
KEIKTFSESGFDGKIVSEVDFGDNKSNEPHEGSDKVTGYTVKLVKVTDDKE
SSVSTTAETKIDLSENKLPVYMHVPKSGALNOKVVFYKGTYPDGSIAGYOWDFD
GSDFSQNSPHVYTKGEYVTLRVMDSSQSEKTMKIKITDPYPIGTEKEPNS
KETASGPIVPGIPVSGIENISQDYFYFDVITPEVKIDINKLGYGAGVAVVYDENN
NAVSAVDDGONLSGKFKADKPGRYIHLVYMFNGSYMPYRINIEGVSGR"
11943..12461
/gene="dcd"
/contig="12461"
/gene="dcd"
/codon_start=1
/transl_table=11
/protein_id="BAA34543.1"
/db_xref="GI:3892647"
/transl_table="MILSGKEIKNRLNKDIFIEPFSNOLNPNNSYNLRLHNLVYEN
NVLDMKENAKKTIPEEGLLLEPGKLYLGRTEHTREKLVPMLEGRSSVGRGLGF
HVIITAGFDGIFSGFWLIEFCVQPIRIYPNIEICQIYYHNIEGEYEKYSTGKYQNT
GQPSLLFKDFE"
12526..13140
/note="orf2d"
/codon_start=1
/transl_table=11
/protein_id="BAA34544.1"
/db_xref="GI:3892648"
/transl_table="MSNKKLLIFIPSFIAIFLLTYGISTFNRNHNKASSNKQGEAK
VFKKHYNLKDINKDQIVESRSINKYQEGKYVIGIKDGSVAIFKVDKNGNMLLESP
KDIITKVEOLPEGDIELLQKDGKLIYENTKEAVEGVEALFRI"
13269..13862
/gene="ruva"
/contig="13862"
/gene="ruva"
/codon_start=1
/transl_table=11
/protein_id="BAA34545.1"
```

```
/db_xref="GI:3892649"

Query Match      0.7%; Score 23; DB 1; Length 14043;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 537 TAAAGACATTAAAGTCTTTCG 559
Db 13760 TAAAGACATTAAAGTCTTTCG 13738

RESULT 33
AC013119/c
LOCUS          AC013119          48626 bp      DNA      linear      HTG 03-NOV-1999
DEFINITION    Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***, in ordered
               pieces.
ACCESSION     AC013119
VERSION       AC013119.1  GI:6223212
KEYWORDS      HTG; HTGS_PHASE2.
SOURCE        Drosophila melanogaster.
ORGANISM      Drosophila melanogaster
               Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
               Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
               Ephydroidea; Drosophilidae; Drosophila.
REFERENCE     1 (bases 1 to 48626)
AUTHORS       Adams,M. and Venter,J.C.
JOURNAL       Direct Submission
               Submitted (03-NOV-1999) Celera Genomics, 45 West Gude Drive,
               Rockville, MD, USA
COMMENT       This sequence was identified as CDM:10213447 by the submitter.
               For further information on this sequence you may e-mail to
               fly@celera.com.
               * NOTE: This is a 'working draft' sequence.
               * This sequence will be replaced
               * by the finished sequence as soon as it is available and
               * the accession number will be preserved.
FEATURES             Location/Qualifiers
     source           1..48626
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
BASE COUNT          13541 a 10753 c 10456 g 13876 t
ORIGIN

Query Match      0.7%; Score 22; DB 2; Length 48626;
Best Local Similarity 100.0%; Pred. No. 7.3;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 266 GGGATGCCAAAATCAATGCCCC 287
Db 31038 GGGATGCCAAAATCAATGCCCC 31017

RESULT 34
DMBR33H2/c
LOCUS          DMBR33H2          85918 bp      DNA      linear      HTG 11-OCT-1999
DEFINITION    Drosophila melanogaster chromosome X clone BACR33H2 map 4B-C strain
               y; cn bw sp, *** SEQUENCING IN PROGRESS ***, in unordered pieces.
ACCESSION     AL121812
VERSION       AL121812.1  GI:6048199
KEYWORDS      HTG; HTGS_PHASE1.
SOURCE        Drosophila melanogaster.
ORGANISM      Drosophila melanogaster
               Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
               Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
               Ephydroidea; Drosophilidae; Drosophila.
REFERENCE     1 (bases 1 to 85918)
AUTHORS       RA Murphy,L., Harris,D. and Barrell,B.
JOURNAL       Sequencing the X chromosome of Drosophila melanogaster
               Unpublished
               Sanger Centre, Wellcome Trust Genome Campus, Hinxton Hall, Hinxton,
               Cambridge CB10 1SA, U.K.
COMMENT       2 (bases 1 to 85918)
REFERENCE     Benos,P.
```

TITLE	Direct Submission	13958	14057: gap of unknown length
JOURNAL	Submitted (10-OCT-1999) European Drosophila Genome Sequencing Consortium	14058	14532: contig of 475 in length
COMMENT	This is a 'working draft' sequence. It currently consists of 147 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence. 1 62: contig of 62 in length	14533	14632: gap of unknown length
		14633	15119: contig of 487 in length
		15120	15219: gap of unknown length
		15220	15725: contig of 506 in length
		15726	15825: gap of unknown length
		15826	16252: contig of 427 in length
		16253	16352: gap of unknown length
		16353	16376: contig of 24 in length
		16377	16476: gap of unknown length
		16477	16816: contig of 340 in length
		16817	16916: gap of unknown length
		16917	17261: contig of 345 in length
		17262	17361: gap of unknown length
		17362	17703: contig of 342 in length
		17704	17803: gap of unknown length
		17804	18078: contig of 275 in length
		18079	18178: gap of unknown length
		18179	19566: contig of 1488 in length
		19667	19756: gap of unknown length
		19677	19857: contig of 91 in length
		19858	19957: gap of unknown length
		19958	19996: contig of 39 in length
		19997	20096: gap of unknown length
		20097	20399: contig of 303 in length
		20400	20499: gap of unknown length
		20500	20940: contig of 441 in length
		20941	21040: gap of unknown length
		21041	21544: contig of 504 in length
		21545	21644: gap of unknown length
		21645	22022: contig of 378 in length
		22023	22122: gap of unknown length
		22123	22429: contig of 307 in length
		22430	22529: gap of unknown length
		22530	23046: contig of 517 in length
		23047	23146: gap of unknown length
		23147	23604: contig of 458 in length
		23605	23704: gap of unknown length
		23705	24181: contig of 477 in length
		24182	24281: gap of unknown length
		24282	24710: contig of 429 in length
		24711	24810: gap of unknown length
		24811	25023: contig of 213 in length
		25024	25123: gap of unknown length
		25124	25580: contig of 457 in length
		25581	25680: gap of unknown length
		25681	25969: contig of 289 in length
		25970	26069: gap of unknown length
		26070	26501: contig of 432 in length
		26502	26601: gap of unknown length
		26602	26908: contig of 307 in length
		26909	27008: gap of unknown length
		27009	27056: contig of 48 in length
		27057	27156: gap of unknown length
		27157	27446: contig of 290 in length
		27447	27546: gap of unknown length
		27547	27740: contig of 194 in length
		27741	27840: gap of unknown length
		27841	28149: contig of 309 in length
		28150	28249: gap of unknown length
		28250	28576: contig of 327 in length
		28577	28676: gap of unknown length
		28677	29126: contig of 450 in length
		29127	29226: gap of unknown length
		29227	29984: contig of 758 in length
		29985	30084: gap of unknown length
		30085	30513: contig of 429 in length
		30514	30613: gap of unknown length
		30614	31009: contig of 396 in length
		31010	31109: gap of unknown length
		31110	31486: contig of 377 in length
		31487	31586: gap of unknown length

31587 31864: contig of 278 in length  
 31865 31964: gap of unknown length  
 31965 32519: contig of 555 in length  
 32520 32619: gap of unknown length  
 32620 32727: contig of 658 in length  
 32728 33377: gap of unknown length  
 33378 33681: contig of 304 in length  
 33682 33782: gap of unknown length  
 33783 34039: contig of 258 in length  
 34040 34139: gap of unknown length  
 34140 34535: contig of 396 in length  
 34536 34635: gap of unknown length  
 34636 35081: contig of 446 in length  
 35082 35181: gap of unknown length  
 35182 35711: contig of 690 in length  
 35712 35971: gap of unknown length  
 35972 36469: contig of 498 in length  
 36470 36569: gap of unknown length  
 36570 37081: contig of 512 in length  
 37082 37181: gap of unknown length  
 37182 37644: contig of 463 in length  
 37645 38170: contig of 426 in length  
 38171 38270: gap of unknown length  
 38271 38821: contig of 551 in length  
 38822 38921: gap of unknown length  
 38922 39658: contig of 737 in length  
 39659 39758: gap of unknown length  
 39759 40361: contig of 603 in length  
 40362 40461: gap of unknown length  
 40462 40942: contig of 481 in length  
 40943 41043: gap of unknown length  
 41043 41429: contig of 387 in length  
 41430 41529: gap of unknown length  
 41530 42050: contig of 521 in length

Query Match 0.7%; Score 22; DB 2; Length 85918;  
 Best Local Similarity 100.0%; Pred. No. 7.4;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 266 GGGATGCCAAATCAATGCCCC 287  
 |||||  
 Db 60216 GGGATGCCAAATCAATGCCCC 60195

RESULT 35  
 AC108481  
 LOCUS  
 DEFINITION Drosophila melanogaster 3L BAC RP98-1019 (Roswell Park Cancer Institute Drosophila BAC Library) complete sequence.  
 ACCESSION AC108481  
 VERSION  
 KEYWORDS HTG.  
 SOURCE fruit fly.  
 ORGANISM Drosophila melanogaster  
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
 REFERENCE 1 (bases 1 to 160986)  
 Gucyney, D., Scherer, S., Adams, M.D., Holt, R.A., Evans, C.A., Muzny, D., Tabor, P., Williamson, A., Homs, F.H., Dugan-Rocha, S.D., Sodergren, E.S., Hodgson, A.H., Chen, R.C., Ayele, M., Scott, G.S., Worley, K.W., Amaratunga, P.G., Brand, R., Rogers, Y., An, H., Baldwin, D., Beeson, K.Y., Brown, M., Buhay, C., Busam, D.A., Center, A., Chen, G., Chen, Z., Clerc-Blankenburg, K., Davenport, L.B., Dietz, S.M., Ding, Y., Dodson, K., Doup, L.E., Draper, H., Emery-Cohen, A., Ferrera, S., Garg, N.D.S., Houck, J., Hostin, D., Howland, T.J., Hume, J., Ibegwam, C., Jalali, M., Kovar, C., Liu, W., Mattel, B., McIntosh, T.C., Morgan, M., Moy, M., Murphy, B., Nelson, K.A., Ndassa, Y., Nguyen, N., Perez, L., Pittman, G.S., Puri, V., Scheeler, F., Shen, H., Strong, R., Tector, C., Wang, Q., Williams, S.M., Xiang, J., Zaveri, J.S., Zhou, J., Zorrilla, S., Smith, H.O., Wheeler, D., Weinstock, G., Gibbs, R. and Venter, J.C.

# TITLE JOURNAL REFERENCE AUTHORS

Direct Submission  
 Unpublished  
 2 (bases 1 to 160986)  
 Worley, K.C., Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C., Alsbrooks, S.L., Amaratunga, H.C., Are, J.R., Banks, T., Barbara, J., Benton, J., Bimage, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyne, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, N., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B., Honsi, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulsegied, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okwuonu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Savary, G., Scherer, S., Scott, G., Shen, H., Shoostari, N., Sisson, I., Sodergren, E., Sonaik, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstock, G. and Gibbs, R.

Direct Submission  
 Unpublished  
 3 (bases 1 to 160986)  
 BCM-HGSC  
 Direct Submission  
 Submitted (29-JAN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
 4 (bases 1 to 160986)  
 BCM-HGSC  
 Direct Submission  
 Submitted (12-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
 INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email [gc-help@bcm.tmc.edu](mailto:gc-help@bcm.tmc.edu)

# TITLE JOURNAL REFERENCE AUTHORS

Direct Submission  
 Unpublished  
 3 (bases 1 to 160986)  
 BCM-HGSC  
 Direct Submission  
 Submitted (29-JAN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
 4 (bases 1 to 160986)  
 BCM-HGSC  
 Direct Submission  
 Submitted (12-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
 INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email [gc-help@bcm.tmc.edu](mailto:gc-help@bcm.tmc.edu)

# COMMENT

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

# ANNOTATION OF FEATURES:

STSS are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.  
 Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

```

FEATURES
    source                1..160986
                        /organism="Drosophila melanogaster"
                        /db_xref="taxon:7227"
                        /chromosome="3L"
                        /clone="RP98-1019"
                        /csize="56335"
                        /function="single clone coverage"
                        /note="one celera read calls extra 5 bps: CGGAC"
                        /function="polymorphic site"
                        /note="size confirmed by PCR and restriction map"
                        /function="unresolved homopolymeric run"
                        /note="size confirmed by PCR and restriction map"
                        /function="unresolved homopolymeric run"

BASE COUNT  46643 a 34050 c 34135 g 46158 t

Query Match      0.7%; Score 22; DB 3; Length 160986;
Best Local Similarity 100.0%; Pred. No. 7.5;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  266  GGGATGCCAAATCAATGCCCC 287
      |||||
Db  22969 GGGATGCCAAATCAATGCCCC 22990

RESULT 36
AC099761
LOCUS      267113 bp  DNA  linear  HTG 23-MAY-2002
DEFINITION Mus musculus chromosome UNK clone RP23-230J6, WORKING DRAFT
SEQUENCE  65 unordered pieces.
AC099761
AC099761.2 GI:21105132
VERSION    HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS   Mus musculus.
SOURCE     Mus musculus
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE  1 (bases 1 to 267113)
AUTHORS   McPherson, J.D. and Waterston, R.H.
TITLE     The sequence of Mus musculus clone
JOURNAL   Unpublished
REFERENCE  2 (bases 1 to 267113)
AUTHORS   McPherson, J.D. and Waterston, R.H.
TITLE     Direct Submission
JOURNAL   Submitted (20-NOV-2001) Genome Sequencing Center, 4444 Forest Park
           Parkway, St. Louis, MO 63108, USA
REFERENCE  3 (bases 1 to 267113)
AUTHORS   McPherson, J.D. and Waterston, R.H.
TITLE     Direct Submission
JOURNAL   Submitted (23-MAY-2002) Genome Sequencing Center, 4444 Forest Park
           Parkway, St. Louis, MO 63108, USA
COMMENT   On May 23, 2002 this sequence version replaced gi:17017633.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@wustl.wustl.edu
----- Project Information -----
Center project name: M_BA0230J06
----- Summary Statistics -----
Sequencing vector: M13; 0%
Chemistry: Dye-primer ET; 0% of reads
Assembly: Dye-terminator Big Dye; 100% of reads
Consensus quality: 241118 bases at least Q40
Consensus quality: 248108 bases at least Q30
Consensus quality: 253552 bases at least Q20
Insert size: 204000; agarose-fp
Insert size: 262842; sum-of-contigs
Quality coverage: 5.94 in Q20 bases; agarose-fp
Quality coverage: 4.19 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 65 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1229: contig of 1229 bp in length
* 1320: gap of unknown length
* 1330: contig of 1226 bp in length
* 2555: gap of unknown length
* 2656: contig of 1121 bp in length
* 3776: gap of unknown length
* 3877: gap of unknown length
* 5042: contig of 1166 bp in length
* 5142: gap of unknown length
* 6206: contig of 1064 bp in length
* 6306: gap of unknown length
* 7391: contig of 1085 bp in length
* 7491: gap of unknown length
* 8871: contig of 1380 bp in length
* 8872: gap of unknown length
* 8971: gap of unknown length
* 10079: contig of 1108 bp in length
* 10080: gap of unknown length
* 10180: contig of 1207 bp in length
* 11387: gap of unknown length
* 11487: contig of 1480 bp in length
* 12967: gap of unknown length
* 13067: contig of 1372 bp in length
* 14439: gap of unknown length
* 14539: contig of 1476 bp in length
* 16015: gap of unknown length
* 16115: contig of 1334 bp in length
* 17449: gap of unknown length
* 17449: gap of unknown length
* 17449: gap of unknown length
* 18774: contig of 1226 bp in length
* 18775: gap of unknown length
* 20311: contig of 1437 bp in length
* 20312: gap of unknown length
* 20412: contig of 1269 bp in length
* 21681: gap of unknown length
* 21781: contig of 1585 bp in length
* 23366: gap of unknown length
* 23466: contig of 1361 bp in length
* 24827: gap of unknown length
* 24927: contig of 1416 bp in length
* 26343: gap of unknown length
* 26443: contig of 1224 bp in length
* 27667: gap of unknown length
* 27667: contig of 1387 bp in length
* 29154: gap of unknown length

```

```
* 29254 30291: contig of 1038 bp in length
* 30292 30391: gap of unknown length
* 30393 31735: contig of 1344 bp in length
* 31736 31835: gap of unknown length
* 31836 32976: contig of 1141 bp in length
* 32977 33076: gap of unknown length
* 33077 34547: contig of 1471 bp in length
* 34548 34647: gap of unknown length
* 34648 36327: contig of 1680 bp in length
* 36328 36427: gap of unknown length
* 36428 37876: contig of 1449 bp in length
* 37877 37976: gap of unknown length
* 37977 39593: contig of 1617 bp in length
* 39594 39693: gap of unknown length
* 39694 40839: contig of 1146 bp in length
* 40840 40939: gap of unknown length
* 40940 43016: contig of 2077 bp in length
* 43017 43116: gap of unknown length
* 43117 44627: contig of 1511 bp in length
* 44628 44727: gap of unknown length
* 44728 46320: contig of 1493 bp in length
* 46321 47884: contig of 1364 bp in length
* 47885 47784: gap of unknown length
* 47785 49420: contig of 1636 bp in length
* 49421 49520: gap of unknown length
* 49521 50957: contig of 1437 bp in length
* 50958 51057: gap of unknown length
* 51058 52514: contig of 1457 bp in length
* 52515 52614: gap of unknown length
* 52615 54905: contig of 2291 bp in length
* 54906 55005: gap of unknown length
* 55006 56149: contig of 1144 bp in length
* 56150 56249: gap of unknown length
* 56250 57770: contig of 1521 bp in length
* 57771 57870: gap of unknown length
* 57871 59283: contig of 1413 bp in length
* 59284 59383: gap of unknown length
* 59384 60656: contig of 1273 bp in length
* 60657 60756: gap of unknown length
* 60757 61903: contig of 1147 bp in length
* 61904 62003: gap of unknown length
* 62004 64818: contig of 2815 bp in length
* 64819 64918: gap of unknown length
* 64919 66167: contig of 1249 bp in length
* 66168 66267: gap of unknown length
* 66268 67997: contig of 1730 bp in length
* 67998 68097: gap of unknown length
* 68098 69365: contig of 1268 bp in length
* 69366 69465: gap of unknown length
* 69466 70770: contig of 1305 bp in length
* 70771 70870: gap of unknown length
* 70871 72937: contig of 2067 bp in length
* 72938 73037: gap of unknown length
* 73038 75049: contig of 2012 bp in length
* 75050 75149: gap of unknown length
* 75150 77650: contig of 2501 bp in length
* 77651 77750: gap of unknown length
* 77751 79482: contig of 1732 bp in length
* 79483 79582: gap of unknown length
* 79583 83087: contig of 3505 bp in length
* 83088 83187: gap of unknown length
* 83188 87699: contig of 4512 bp in length
* 87700 87799: gap of unknown length
* 87800 93702: contig of 5903 bp in length
* 93703 93802: gap of unknown length
* 93803 99216: contig of 5414 bp in length
* 99217 99316: gap of unknown length
* 99317 105763: contig of 6447 bp in length
* 105764 105864: gap of unknown length
* 105865 113706: contig of 7843 bp in length
* 113707 113806: gap of unknown length
* 113807 122387: contig of 8581 bp in length
```

```
* 122388 122487: gap of unknown length
* 122488 133994: contig of 11407 bp in length
* 133995 133995: gap of unknown length
* 133995 143826: contig of 9832 bp in length
* 143827 143926: gap of unknown length
* 143927 162163: contig of 18237 bp in length
* 162164 162263: gap of unknown length
* 162264 179812: contig of 17549 bp in length
* 179813 179912: gap of unknown length
* 179913 200350: contig of 20438 bp in length
* 200351 200450: gap of unknown length
* 200451 232908: contig of 32458 bp in length
* 232909 233008: gap of unknown length
* 233009 267113: contig of 34105 bp in length.
FEATURES             Location/Qualifiers
     source            1..267113
                       /organism="Mus musculus"
                       /db_xref="taxon:10090"
                       /chromosome="UNK"
                       /clone="RP23-230J6"
     misc_feature      1..1229
                       /note="assembly_name:Contig35"
     misc_feature      1330..2555
                       /note="assembly_name:Contig45"
     misc_feature      2656..3776
                       /note="assembly_name:Contig55"
     misc_feature      3877..5042
                       /note="assembly_name:Contig95"
     misc_feature      5143..6206

Query Match          0.7%; Score 22; DB 2; Length 267113;
Best Local Similarity 100.0%; Pred. No. 7.6;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1382 TGGATCGCGGTAAAGCCATGAA 1403
      |||||
DB,182950 TGGATCGCGGTAAAGCCATGAA 182971

RESULT 37
AE003430
LOCUS      294218 bp      DNA      linear      INV 04-OCT-2000
DEFINITION Drosophila melanogaster genomic scaffold 142000013386054 section 14
            of 35, complete sequence.
ACCESSION  AE003430 AE002566
VERSION     AE003430.2 GI:10728458
KEYWORDS    HTG.
SOURCE      Drosophila melanogaster.
ORGANISM    Drosophila melanogaster.
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 294218)
AUTHORS     Adams,M.D., Celnikier,S.E., Holt,R.A., Evans,C.A., Gocayne,J.D.,
            Amanatides,P.G., Scherer,S.E., Li,P.W., Hoskins,R.A., Galle,R.F.,
            George,R.A., Lewis,S.E., Richards,S., Ashburner,M., Henderson,S.N.,
            Sutton,G.G., Wortman,J.R., Vandeil,M.D., Zhang,Q., Chen,L.X.,
            Brandon,R.C., Rogers,Y.H., Blazej,R.G., Champe,M., Pfeiffer,B.D.,
            Wan,K.H., Doyle,C., Baxter,E.G., Helt,G., Nelson,C.R., Gabor
            Miklos,G.L., Abril,J.F., Agbayani,A., An,H.J.,
            Andrews-Pfannkoch,C., Baldwin,D., Ballew,R.M., Basu,A.,
            Baxendale,J., Bayraktaroglu,L., Beasley,E.M., Beeson,K.Y.,
            Benos,P.V., Berman,B.P., Bhandari,D., Bolshakov,S., Borkova,D.,
            Bortchan,M.R., Bouck,J., Brokstein,P., Brottier,P., Burtis,K.C.,
            Busam,D.A., Butler,H., Cadieu,E., Center,A., Chandra,I.,
            Cherry,J.M., Cui,Y., Dahlke,C., Davenport,L.B., Davies,P., de
            Pablos,B., Delcher,A., Deng,Z., Mays,A.D., Dew,I., Dietz,S.M.,
            Dodson,K., Doup,L.E., Downes,M., Dugan-Rocha,S., Dunkov,B.C.,
            Dunn,P., Durbin,K.J., Evangelista,C.C., Ferraz,C., Ferreira,S.,
            Fleischmann,W., Foster,C., Gabrielian,A.E., Garg,N.S.,
            Gelbart,W.M., Glasser,K., Glodek,A., Gong,F., Gorrell,J.H., Gu,Z.,
            Guan,P., Harris,M., Harris,N.L., Harvey,D., Heiman,T.J.,
            Hernandez,J.R., Houck,J., Hostin,D., Houston,K.A., Howland,T.J.,
```





```

/ gene="EG:EG0007.11"
/ note="CG4944"
/ map="4B1-4B1"
/ db_xref="FLYBASE:FBan0004944"
/ db_xref="FLYBASE:FBgn0026084"
join(11171..11855,15271..15406,15479..15592,15734..16524)
/ gene="EG:EG0007.11"
/ product="CT15683"
/ db_xref="FLYBASE:FBan0004944"
/ db_xref="FLYBASE:FBgn0026084"
join(12417..12509,15271..15406,15479..15592,15734..16524)
/ gene="EG:EG0007.11"
/ product="CT15665"
/ db_xref="FLYBASE:FBan0004944"
/ db_xref="FLYBASE:FBgn0026084"
join(15271..15406,15479..15592,15734..15873)
/ gene="EG:EG0007.11" gene product {alt 1}
/ note="EG:EG0007.11"
/ codon_start=1
/ protein_id="AAF45919.1"
/ db_xref="GI:7290465"
/ db_xref="FLYBASE:FBan0004944"

Query Match      0.7%; Score 22; DB 3; Length 294218;
Best Local Similarity 100.0%; Pred. No. 7.6;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 266 GGGATGCCAAATCAATGCCCC 287
      |||||
Db 158358 GGGATGCCAAATCAATGCCCC 158379

RESULT 38
AX024141
LOCUS AX024141 34 bp DNA linear PAT 15-SEP-2000
DEFINITION Sequence 84 from Patent FR2785293.
ACCESSION AX024141
VERSION AX024141.1 GI:10184452
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 34)
AUTHORS Nassif,X., Tinsley,C., Aujame,L., Perrin,A., Rokbi,B.,
Bouchardon,A. and Renaud,M.G.
JOURNAL Patent: FR 2785293-A 84 05-MAY-2000;
PASTEUR MERIEUX SERUMS VACC (FR)
FEATURES
    source
        1..34
        /organism="synthetic construct"
        /db_xref="taxon:32630"
        /note="amorce de"
BASE COUNT 9 a 15 c 5 g 5 t
ORIGIN

Query Match      0.7%; Score 21; DB 6; Length 34;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCGAACGACCCCAACCTTC 21
      |||||
Db 14 ATGCGAACGACCCCAACCTTC 34

RESULT 39
AF281339/c
LOCUS AF281339 348 bp mRNA linear VRT 20-FEB-2002
DEFINITION Oncorhynchus mykiss hemopexin-like protein variant 1 mRNA, partial cds.
ACCESSION AF281339
VERSION AF281339.1 GI:11095770
KEYWORDS
SOURCE Oncorhynchus mykiss.

Oncorhynchus mykiss
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
1 (bases 1 to 348)
Bayne,C.J., Gerwick,L., Fujiki,K., Nakao,M. and Yano,T.
Immune-relevant (including acute phase) genes identified in the
livers of rainbow trout, Oncorhynchus mykiss, by means of
suppression subtractive hybridization
Dev. Comp. Immunol. 25 (3), 205-217 (2001)
21098507
MEDLINE 11164886
PUBMED
REFERENCE 2 (bases 1 to 348)
AUTHORS Bayne,C.J., Gerwick,L., Fujiki,K., Nakao,M. and Yano,T.
TITLE Direct Submission
JOURNAL Submitted (20-JUN-2000) Zoology, Oregon State University, 3029
Cordley, Corvallis, OR 97331, USA
FEATURES
    source
        1..348
        /organism="Oncorhynchus mykiss"
        /db_xref="taxon:8022"
        <1..>348
        /function="binds free heme"
        /codon_start=1
        /product="hemopexin-like protein variant 1"
        /protein_id="AAG30014.1"
        /db_xref="GI:11095771"
        /translation="NGHTFFKGDHLNFKGPAQVSSAFKELDNYHHLGHVDAAFR
MHNREKPEKHDIIFLDKRVFSYINHSLEEGYPDIQIDFGPFSHVDAAVECPKGE
CNSDSVLFFKGEV"
BASE COUNT 91 a 94 c 87 g 74 t 2 others
ORIGIN

Query Match      0.7%; Score 21; DB 5; Length 348;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2710 CATGCGAAGGAGCGCTCAAC 2730
      |||||
Db 135 CATGCGAAGGAGCGCTCAAC 115

RESULT 40
HSA332365/c
LOCUS HSA332365 768 bp DNA linear PRI 18-JUL-2002
DEFINITION Homo sapiens genomic sequence surrounding NotI site, clone
NR3-FE23RS.
ACCESSION AJ332365
VERSION AJ332365.1 GI:15876783
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 768)
Kutsenko,A.S., Gizatullin,R.Z., Al-Amin,A.N., Wang,F., Kvasha,S.M.,
Podowski,R.M., Matushkin,Y.G., Gvanchandani,A., Muravenko,O.V.,
Levitsky,V.G., Kolchanov,N.A., Protodopov,A.I., Kashuba,V.I.,
Kisselev,L.L., Wasserman,W., Wahlestedt,C. and Zabarovsky,E.R.
NotI flanking sequences: a tool for gene discovery and verification
of the human genome
Nucleic Acids Res. 30 (14), 3163-3170 (2002)
12136098
PUBMED
REFERENCE 2 (bases 1 to 768)
AUTHORS Zabarovsky,E.R.
TITLE Direct Submission
JOURNAL Submitted (16-MAY-2001) Microbiology and Tumorbiology Centre,
Karolinska Institute, Theorells vag, 3, Box 280, Stockholm 171 77,
Sweden
FEATURES
    source
        Location/Qualifiers
        1..768
        /organism="Homo sapiens"
        /db_xref="taxon:9606"

```

BASE COUNT 185 a 205 c 204 g 174 t  
 ORIGIN /clone="NR5-FE23RS"

Query Match 0.78; Score 21; DB 9; Length 768;  
 Best Local Similarity 100.0%; Pred. No. 24;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2420 AAACCCAAACAGGCGGTGGAA 2440  
 ||||||||||||||||||  
 Db 537 AAACCCAAACAGGCGGTGGAA 517

Search completed: January 27, 2003, 23:33:45  
 Job time : 10122 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 27, 2003, 20:44:54 ; Search time 2874 Seconds  
(without alignments)  
18055.094 Million cell updates/sec

Title: US-09-830-433A-7  
Perfect score: 3204  
Sequence: 1 atcggaacgacccaactt.....gcgtaggctaccggttctga 3204

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database : EST:\*

1: em\_estba:\*  
2: em\_estbam:\*  
3: em\_estin:\*  
4: em\_estimu:\*  
5: em\_estov:\*  
6: em\_estovl:\*  
7: em\_estro:\*  
8: em\_estro:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_est3:\*  
12: gb\_est4:\*  
13: gb\_est5:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	25	0.8	123	17	A0024974
C 2	25	0.8	365	17	BH759360
C 3	25	0.8	431	13	BI241907
C 4	25	0.8	440	13	BI482365
C 5	25	0.8	443	9	AI063687
C 6	25	0.8	485	13	BI365464
C 7	25	0.8	485	13	BI365464
C 8	25	0.8	485	13	BI365464
C 9	25	0.8	485	13	BI365464
C 10	25	0.8	485	13	BI365464
C 11	25	0.8	485	13	BI365464
C 12	25	0.8	485	13	BI365464
C 13	25	0.8	485	13	BI365464
C 14	25	0.8	485	13	BI365464
C 15	25	0.8	485	13	BI365464
C 16	25	0.8	485	13	BI365464
C 17	25	0.8	485	13	BI365464
C 18	25	0.8	485	13	BI365464
C 19	25	0.8	485	13	BI365464
C 20	25	0.8	485	13	BI365464
C 21	25	0.8	485	13	BI365464
C 22	25	0.8	485	13	BI365464
C 23	25	0.8	485	13	BI365464
C 24	25	0.8	485	13	BI365464
C 25	25	0.8	485	13	BI365464
C 26	25	0.8	485	13	BI365464
C 27	25	0.8	485	13	BI365464
C 28	25	0.8	485	13	BI365464
C 29	25	0.8	485	13	BI365464
C 30	25	0.8	485	13	BI365464
C 31	25	0.8	485	13	BI365464
C 32	25	0.8	485	13	BI365464
C 33	25	0.8	485	13	BI365464
C 34	25	0.8	485	13	BI365464
C 35	25	0.8	485	13	BI365464
C 36	25	0.8	485	13	BI365464
C 37	25	0.8	485	13	BI365464
C 38	25	0.8	485	13	BI365464
C 39	25	0.8	485	13	BI365464
C 40	25	0.8	485	13	BI365464
C 41	25	0.8	485	13	BI365464
C 42	25	0.8	485	13	BI365464
C 43	25	0.8	485	13	BI365464
C 44	25	0.8	485	13	BI365464
C 45	25	0.8	485	13	BI365464
C 46	25	0.8	485	13	BI365464
C 47	25	0.8	485	13	BI365464
C 48	25	0.8	485	13	BI365464
C 49	25	0.8	485	13	BI365464
C 50	25	0.8	485	13	BI365464
C 51	25	0.8	485	13	BI365464
C 52	25	0.8	485	13	BI365464
C 53	25	0.8	485	13	BI365464
C 54	25	0.8	485	13	BI365464
C 55	25	0.8	485	13	BI365464
C 56	25	0.8	485	13	BI365464
C 57	25	0.8	485	13	BI365464
C 58	25	0.8	485	13	BI365464
C 59	25	0.8	485	13	BI365464
C 60	25	0.8	485	13	BI365464
C 61	25	0.8	485	13	BI365464
C 62	25	0.8	485	13	BI365464
C 63	25	0.8	485	13	BI365464
C 64	25	0.8	485	13	BI365464
C 65	25	0.8	485	13	BI365464
C 66	25	0.8	485	13	BI365464
C 67	25	0.8	485	13	BI365464
C 68	25	0.8	485	13	BI365464
C 69	25	0.8	485	13	BI365464
C 70	25	0.8	485	13	BI365464
C 71	25	0.8	485	13	BI365464
C 72	25	0.8	485	13	BI365464
C 73	25	0.8	485	13	BI365464
C 74	25	0.8	485	13	BI365464
C 75	25	0.8	485	13	BI365464
C 76	25	0.8	485	13	BI365464
C 77	25	0.8	485	13	BI365464
C 78	25	0.8	485	13	BI365464
C 79	25	0.8	485	13	BI365464
C 80	25	0.8	485	13	BI365464
C 81	25	0.8	485	13	BI365464
C 82	25	0.8	485	13	BI365464
C 83	25	0.8	485	13	BI365464
C 84	25	0.8	485	13	BI365464
C 85	25	0.8	485	13	BI365464
C 86	25	0.8	485	13	BI365464
C 87	25	0.8	485	13	BI365464
C 88	25	0.8	485	13	BI365464
C 89	25	0.8	485	13	BI365464
C 90	25	0.8	485	13	BI365464
C 91	25	0.8	485	13	BI365464
C 92	25	0.8	485	13	BI365464
C 93	25	0.8	485	13	BI365464
C 94	25	0.8	485	13	BI365464
C 95	25	0.8	485	13	BI365464
C 96	25	0.8	485	13	BI365464
C 97	25	0.8	485	13	BI365464
C 98	25	0.8	485	13	BI365464
C 99	25	0.8	485	13	BI365464
C 100	25	0.8	485	13	BI365464

7	25	0.8	491	13	BI165260
8	25	0.8	521	13	BI227520
9	25	0.8	526	13	BI363907
10	25	0.8	531	13	BI356933
11	25	0.8	544	9	AI114180
12	25	0.8	545	9	AI135320
13	25	0.8	560	13	BI362358
14	25	0.8	577	13	BI164055
15	25	0.8	579	13	BI165486
16	25	0.8	583	17	AQ074074
17	25	0.8	586	17	AQ073492
18	25	0.8	590	9	AI387965
19	25	0.8	590	13	BI238617
20	25	0.8	592	13	BI363426
21	25	0.8	593	13	BI481998
22	25	0.8	602	9	AI109244
23	25	0.8	616	9	AI107819
24	25	0.8	625	9	AI063205
25	25	0.8	625	9	AI238498
26	25	0.8	626	13	BI170514
27	25	0.8	628	9	AI296017
28	25	0.8	633	13	BI169402
29	25	0.8	639	9	AI386994
30	25	0.8	655	9	AI388097
31	25	0.8	664	9	AI107422
32	25	0.8	664	9	AI292862
33	25	0.8	666	9	AI109291
34	25	0.8	677	13	BI368317
35	25	0.8	686	13	BI374631
36	25	0.8	687	13	BI368263
37	25	0.8	734	9	AI517363
38	21	0.7	496	17	AZ628307
39	21	0.7	593	13	BI468065
40	21	0.7	727	9	AI238153
41	21	0.7	762	12	BF494167
42	21	0.7	803	17	CNS05E0W
43	21	0.7	1010	17	CNS03XQ0
44	21	0.7	1068	17	CNS058XF
45	21	0.6	182	17	AZ577491
46	20	0.6	244	13	BI321685
47	20	0.6	345	12	BG588535
48	20	0.6	375	12	AZ470268
49	20	0.6	464	14	BQ450589
50	20	0.6	504	9	AL817975
51	20	0.6	559	17	TA27E05Q
52	20	0.6	579	13	BM279219
53	20	0.6	597	17	AZ214569
54	20	0.6	598	17	AQ941303
55	20	0.6	608	13	BI980081
56	20	0.6	658	17	BI189926
57	20	0.6	756	13	BI246500
58	20	0.6	796	13	BM289442
59	20	0.6	967	12	BF141559
60	20	0.6	1146	14	BQ734165
61	19	0.6	211	12	BF758324
62	19	0.6	213	9	AA040282
63	19	0.6	222	17	B79680
64	19	0.6	255	14	R85275
65	19	0.6	268	10	BB563246
66	19	0.6	284	9	AA172297
67	19	0.6	287	17	BQ571703
68	19	0.6	305	13	BM429089
69	19	0.6	313	9	AA046507
70	19	0.6	315	9	AI767192
71	19	0.6	346	9	AI478550
72	19	0.6	364	12	BG878645
73	19	0.6	368	10	AW016361
74	19	0.6	379	10	AW016352
75	19	0.6	380	10	BB841859
76	19	0.6	381	9	AA128881
77	19	0.6	396	9	AA680349
78	19	0.6	401	10	AW752006
79	19	0.6	404	17	AQ445088

80	19	0.6	412	9	AI767499	AI767499 wh27d01.x	153	19	0.6	737	12	BF796989	BF796989 602258020		
81	19	0.6	418	17	AQ0877258	AQ0877258 HS_2142_B	c 154	19	0.6	740	17	BH838125	BH838125 LMC910000		
82	19	0.6	421	17	AQ088119	AQ088119 HS_2191_B	c 155	19	0.6	745	14	BU005179	BU005179 QCG7F17.Y		
83	19	0.6	426	14	BM712611	BM712611 UI-E-EJ0-	c 156	19	0.6	748	13	BI087971	BI087971 602852435		
84	19	0.6	427	17	AQ966628	AQ966628 LERILL6TR	c 157	19	0.6	749	9	AJ448695	AJ448695 AJ448695		
85	19	0.6	429	10	AW469680	AW469680 h433a04.x	c 158	19	0.6	756	9	AJ451784	AJ451784 AJ451784		
86	19	0.6	448	9	AA047592	AA047592 zfi3e03.r	c 159	19	0.6	759	14	BQ863081	BQ863081 QGC22018.		
87	19	0.6	457	12	BE992116	BE992116 UI-M-BZ1-	c 160	19	0.6	767	13	BI857651	BI857651 603384966		
88	19	0.6	473	9	AI492996	AI492996 qz47e08.x	c 161	19	0.6	771	12	BE796734	BE796734 601587357		
89	19	0.6	477	10	BB852900	BB852900 BB852900	c 162	19	0.6	789	10	BE544586	BE544586 601077211		
90	19	0.6	500	9	AL568271	AL568271 AL568271	c 163	19	0.6	789	12	BG696610	BG696610 602658907		
91	19	0.6	506	12	BG482259	BG482259 602526782	c 164	19	0.6	793	17	BH034274	BH034274 RPCI-24-2		
92	19	0.6	513	17	AQ141768	AQ141768 HS_3154_B	c 165	19	0.6	803	12	BG696444	BG696444 602659552		
93	19	0.6	517	17	AQ445307	AQ445307 GSSTC0159	c 166	19	0.6	836	11	AV104690	AV104690 Zee may's		
94	19	0.6	522	17	AQ049847	AQ049847 LNAJFV1_1	c 167	19	0.6	839	12	BF621570	BF621570 HVSMEa001		
95	19	0.6	528	12	BF540443	BF540443 602052414	c 168	19	0.6	843	17	AQ889018	AQ889018 HS_3144_A		
96	19	0.6	530	17	AQ675820	AQ675820 HS_2145_A	c 169	19	0.6	850	12	BF308037	BF308037 601889419		
97	19	0.6	533	14	BM888527	BM888527 TM198_Hum	c 170	19	0.6	850	14	BQ425757	BQ425757 AGENCOURT		
98	19	0.6	534	9	AA777559	AA777559 z194d05.s	c 171	19	0.6	852	17	CNS0102P	AL099631 Dtosophi1		
99	19	0.6	537	13	BG930136	BG930136 etSHESF00	c 172	19	0.6	854	14	BQ428170	BQ428170 AGENCOURT		
100	19	0.6	539	9	AI156683	AI156683 z108c06.r	c 173	19	0.6	881	9	AL540182	AL540182 AL540182		
101	19	0.6	544	13	BM3133985	BM3133985 i05d09.Y	c 174	19	0.6	891	14	BQ423933	BQ423933 AGENCOURT		
102	19	0.6	545	13	BM272641	BM272641 i0i0i07.x	c 175	19	0.6	913	9	AL557429	AL557429 AL557429		
103	19	0.6	550	17	AQ578653	AQ578653 nbxb0093C	c 176	19	0.6	914	17	AZ692320	AZ692320 ENTILL9TR		
104	19	0.6	558	14	W67937	W67937 zd38g12.r1	c 177	19	0.6	916	9	AL529762	AL529762 AL529762		
105	19	0.6	562	9	AI033251	AI033251 ow98e07.s	c 178	19	0.6	920	9	AL537237	AL537237 AL537237		
106	19	0.6	563	17	AQ891918	AQ891918 HS_3154_B	c 179	19	0.6	920	9	AL563543	AL563543 AL563543		
107	19	0.6	566	9	AI522084	AI522084 ti83e08.x	c 180	19	0.6	941	12	BE783156	BE783156 601470765		
108	19	0.6	567	10	BE587706	BE587706 WHE00667.C	c 181	19	0.6	955	14	BQ918272	BQ918272 AGENCOURT		
109	19	0.6	567	13	BI771188	BI771188 603054892	c 182	19	0.6	980	9	AL567771	AL567771		
110	19	0.6	568	13	BI726466	BI726466 i031085E0	c 183	19	0.6	988	17	CNS04BT9	AG072512 Pan trogl		
111	19	0.6	570	13	BI779411	BI779411 EBF001.SQ	c 184	19	0.6	998	17	AG072512	AG072512 Pan trogl		
112	19	0.6	573	13	BM526937	BM526937 sal47e03.	c 185	19	0.6	1006	14	BQ944799	BQ944799 AGENCOURT		
113	19	0.6	575	17	AZ378295	AZ378295 iM0133C06	c 186	19	0.6	1011	9	AL542355	AL542355 AL542355		
114	19	0.6	578	14	BU005018	BU005018 QG66N19.Y	c 187	19	0.6	1036	9	AL542354	AL542354 AL542354		
115	19	0.6	579	9	AA626793	AA626793 ad09g12.s	c 188	19	0.6	1048	13	BM451556	BM451556 AGENCOURT		
116	19	0.6	589	14	BQ776091	BQ776091 UI-H-FH0-	c 189	19	0.6	1147	11	AK009713	AK009713 Mus muscu		
117	19	0.6	592	17	BM835120	BM835120 BACP7-F0	c 190	19	0.6	1166	12	BE793668	BE793668 601588747		
118	19	0.6	596	14	BM943333	BM943333 UI-M-EH0P	c 191	19	0.6	1364	13	BM015155	BM015155 603641146		
119	19	0.6	600	13	BI991111	BI991111 4091-20.M	c 192	19	0.6	1474	13	BM473239	BM473239 AGENCOURT		
120	19	0.6	602	9	AJ447794	AJ447794 AJ447794	c 193	19	0.6	1532	17	AG121472	AG121472 Pan trogl		
121	19	0.6	608	17	BH692403	BH692403 BOHZS96TF	c 194	19	0.6	1924	12	BF527558	BF527558 602040481		
122	19	0.6	610	9	AA150480	AA150480 z108c06.s	c 195	19	0.6	1942	13	BM470504	BM470504 AGENCOURT		
123	19	0.6	611	17	AQ271191	AQ271191 nbxb0024K	c 196	19	0.6	18	0.6	118	10	AW708999	AW708999 d3allne.r
124	19	0.6	616	10	BM631390	BM631390 BB631390	c 197	19	0.6	1364	13	BM015155	BM015155 603641146		
125	19	0.6	629	10	AV649619	AV649619 AV649619	c 198	19	0.6	167	10	BB563762	BB563762 BB563762		
126	19	0.6	629	10	AW375504	AW375504 QV0-C3018	c 199	19	0.6	171	13	BI764138	BI764138 603043368		
127	19	0.6	631	17	AZ995907	AZ995907 2M0281H23	c 200	19	0.6	171	17	AZ285464	AZ285464 RPCI-23-4		
128	19	0.6	634	9	AJ448948	AJ448948 AJ448948	c 201	19	0.6	172	13	BI771966	BI771966 603058996		
129	19	0.6	643	9	AJ454502	AJ454502 AJ454502	c 202	19	0.6	181	13	BI824405	BI824405 603038637		
130	19	0.6	645	12	BG414880	BG414880 HVSMEK000	c 203	19	0.6	185	13	BI915403	BI915403 603177519		
131	19	0.6	649	9	AJ451907	AJ451907 AJ451907	c 204	19	0.6	190	10	AW721882	AW721882 alcl2nm.r		
132	19	0.6	655	17	BH374446	BH374446 AG-ND-179	c 205	19	0.6	192	10	AW711311	AW711311 fid07ne.r		
133	19	0.6	656	14	BM990290	BM990290 UI-H-DH0-	c 206	19	0.6	207	9	AI319959	AI319959 b7f02nm.r		
134	19	0.6	657	17	AQ398889	AQ398889 mgxb00110	c 207	19	0.6	211	10	AW879740	AW879740 QV3-OT002		
135	19	0.6	660	14	BQ573802	BQ573802 UI-H-EZ0-	c 208	19	0.6	223	10	BB591304	BB591304 BB591304		
136	19	0.6	661	12	BE834042	BE834042 MR2-OT007	c 209	19	0.6	227	9	AV253227	AV253227 AV253227		
137	19	0.6	663	14	BQ573850	BQ573850 UI-H-EZ0-	c 210	19	0.6	230	10	AW254220	AW254220 UI-R-BT0-		
138	19	0.6	667	14	BQ009318	BQ009318 UI-H-ED1-	c 211	19	0.6	241	10	AW222147	AW222147 EST298958		
139	19	0.6	671	14	BQ574576	BQ574576 UI-H-EZ1-	c 212	19	0.6	246	9	AI319735	AI319735 b5c09nm.r		
140	19	0.6	679	14	BM714805	BM714805 UI-E-EJ0-	c 213	19	0.6	247	12	BG209826	BG209826 RST29354		
141	19	0.6	682	10	AW235106	AW235106 xn18f07.x	c 214	19	0.6	247	13	BG933351	BG933351 WSL_3_E10		
142	19	0.6	688	12	BG479877	BG479877 602527419	c 215	19	0.6	253	17	AZ561628	AZ561628 RPCI-23-1		
143	19	0.6	691	9	AJ446198	AJ446198 AJ446198	c 216	19	0.6	254	10	AW722970	AW722970 c9p12nm.r		
144	19	0.6	694	12	BE731312	BE731312 601568430	c 217	19	0.6	255	12	EG788767	EG788767 SEADMC008		
145	19	0.6	695	14	BQ863615	BQ863615 QGC24F21.	c 218	19	0.6	256	9	AI426086	AI426086 mh36a03.x		
146	19	0.6	698	12	BG847722	BG847722 1024019D0	c 219	19	0.6	258	9	AA715375	AA715375 nv52a08.r		
147	19	0.6	704	14	BM977948	BM977948 UI-CF-EC1	c 220	19	0.6	258	12	BF290318	BF290318 EST454909		
148	19	0.6	710	9	AJ453198	AJ453198 AJ453198	c 221	19	0.6	263	10	AW709370	AW709370 d4h03ne.r		
149	19	0.6	714	14	BQ860993	BQ860993 QGC17D21.	c 222	19	0.6	265	9	AL840777	AL840777 AL840777		
150	19	0.6	720	14	BQ862390	BQ862390 QGC20F11.	c 223	19	0.6	268	10	BE211550	BE211550 so63col.Y		
151	19	0.6	725	9	AJ449948	AJ449948 AJ449948	c 224	19	0.6	271	10	AV629344	AV629344 AV629344		
152	19	0.6	730	9	AJ455348	AJ455348 AJ455348	c 225	19	0.6	273	10	BE425248	BE425248 WHE313_A0		

226	18	0.6	279	10	BB564822	BB564822	BB564822	299	18	0.6	390	10	AW715157	AW715157	g2e12nm..r
227	18	0.6	280	9	AW176829	AW176829	EST20429	300	18	0.6	390	10	AW718187	AW718187	i4b12nm..r
228	18	0.6	283	10	AW229465	AW229465	uo59c06.y	301	18	0.6	390	10	AW722119	AW722119	a4e08nm..r
229	18	0.6	284	10	AW718106	AW718106	i3g02nm..r	302	18	0.6	392	10	AW724846	AW724846	q173c06.x
230	18	0.6	284	10	BB045431	BB045431	BB045431	c 303	18	0.6	392	14	D492436	D492436	RTCS16373A
231	18	0.6	286	17	CNS03TK6	AL259935	Tetraodon	c 304	18	0.6	394	9	AA992621	AA992621	ot81a01.s
232	18	0.6	292	10	AW708555	AW708555	g7dl1ne.f	c 305	18	0.6	394	17	AO489855	AO489855	RPCI-11-2
233	18	0.6	294	9	A1320261	A1320261	c2e12nm..r	c 306	18	0.6	395	9	AI646144	AI646144	uc41e03.x
234	18	0.6	297	13	B1999212	B1999212	1031071G0	c 307	18	0.6	395	14	HA16968	HA16968	yo19q09.r1
235	18	0.6	298	10	BE341778	BE341778	EST394610	c 308	18	0.6	399	9	A1327735	A1327735	10c06a1.f
236	18	0.6	300	14	R53424	R53424	y170h06.r1	c 309	18	0.6	399	12	BG135223	BG135223	EST468115
237	18	0.6	303	12	BF754561	BF754561	OV3-CT056	c 310	18	0.6	400	9	AA618058	AA618058	nq03a12.s
238	18	0.6	305	12	BE950617	BE950617	UI-M-CE0-	c 311	18	0.6	401	9	AA488254	AA488254	ad08f02.r
239	18	0.6	309	9	AA031134	AA031134	mi47b09.r	c 312	18	0.6	401	14	BM824436	BM824436	K-EST0095
240	18	0.6	312	13	B1313612	B1313612	da480f10.	c 313	18	0.6	402	9	AA4848564	AA4848564	EST191324
241	18	0.6	313	9	A1581489	A1581489	ar63d10.x	c 314	18	0.6	402	9	AA848564	AA848564	EST191324
242	18	0.6	314	9	A1639612	AI639612	vv50c02.x	c 315	18	0.6	404	9	AA303191	AA303191	EST15945
243	18	0.6	314	10	AW718452	AW718452	i5d09nm..r	c 316	18	0.6	404	12	BG097096	BG097096	EST461615
244	18	0.6	315	9	AO065636	AO065636	AO065636	c 317	18	0.6	405	9	AA663519	AA663519	ab80c09.s
245	18	0.6	316	14	BM793718	BM793718	K-EST0074	c 318	18	0.6	406	9	AI265847	AI265847	q111e01.x
246	18	0.6	317	17	AZ004714	AZ004714	RPCI-23-3	c 319	18	0.6	409	10	AW406347	AW406347	UI-HF-BL0
247	18	0.6	319	10	BE646187	BE646187	7e82q03.x	c 320	18	0.6	409	12	BG736828	BG736828	rk63a05.y
248	18	0.6	321	10	AW796612	AW796612	CM1-UM003	c 321	18	0.6	411	9	A1352374	A1352374	qt21a04.x
249	18	0.6	326	10	AW139530	AW139530	UI-H-B11-	c 322	18	0.6	411	17	AQ420481	AQ420481	RPCT-11-1
250	18	0.6	326	10	AW711619	AW711619	f4c08ne.r	c 323	18	0.6	413	10	BB688204	BB688204	BB688204
251	18	0.6	327	17	BH233280	BH233280	1006172G0	c 324	18	0.6	414	9	AI700328	AI700328	wd06d04.x
252	18	0.6	328	13	BG950118	BG950118	CM2-CT063	c 325	18	0.6	414	10	AV640008	AV640008	AV640008

c 372	18	0.6	449	9	AT989401	AT989401 wt80g06.x	445	18	0.6	510	13	BM123923	BM123923
c 373	18	0.6	450	9	AT744144	AT744144 wc36g09.x	446	18	0.6	512	17	BE601703	BE601703
c 374	18	0.6	450	9	AA596973	AA596973 vo16g07.r	447	18	0.6	515	10	AQ173539	AQ173539
c 375	18	0.6	450	17	BH233160	BH233160 1006172A0	448	18	0.6	517	12	BG383572	BG383572
c 376	18	0.6	451	9	AT311704	AT311704 uk17g05.y	c 449	18	0.6	517	12	BF257675	BF257675
c 377	18	0.6	451	9	AT1810755	AT1810755 tu04c02.x	450	18	0.6	517	14	BQ241055	BQ241055
c 378	18	0.6	452	9	AT358863	AT358863 qy24a04.x	451	18	0.6	519	10	BE453917	BE453917
c 379	18	0.6	453	10	AV866472	AV866472 AV866472	c 452	18	0.6	520	12	BG131438	BG131438
c 380	18	0.6	455	9	AA848571	AA848571 EST191331	c 453	18	0.6	521	9	AA003883	AA003883
c 381	18	0.6	455	9	AU174161	AU174161 AU174161	c 454	18	0.6	521	9	AI684218	AI684218
c 382	18	0.6	456	10	BB859431	BB859431 BB859431	455	18	0.6	523	17	AQ592249	AQ592249
c 383	18	0.6	457	17	AQ936179	AQ936179 RPCI-11-S	456	18	0.6	524	10	BE415327	BE415327
c 384	18	0.6	458	10	AW049881	AW049881 UI-M-BH1-	457	18	0.6	525	14	BM792906	BM792906
c 385	18	0.6	458	10	BB751397	BB751397 BB751397	458	18	0.6	528	17	BH217120	BH217120
c 386	18	0.6	458	10	BB826918	BB826918 BB826918	c 459	18	0.6	529	10	AW341623	AW341623
c 387	18	0.6	458	10	BE630893	BE630893 uu50e02.x	c 460	18	0.6	529	10	BE445703	BE445703
c 388	18	0.6	459	17	BH624494	BH624494 1007078G0	461	18	0.6	529	12	BG312601	BG312601
c 389	18	0.6	460	9	AA309020	AA309020 EST179803	c 462	18	0.6	531	12	BG135151	BG135151
c 390	18	0.6	461	10	AW279182	AW279182 sf67e08.y	c 463	18	0.6	532	10	BE444831	BE444831
c 391	18	0.6	461	14	BM852619	BM852619 K-EST0133	c 464	18	0.6	532	10	BE471739	BE471739
c 392	18	0.6	462	9	AL527075	AL527075 AL527075	c 465	18	0.6	533	12	BF098566	BF098566
c 393	18	0.6	462	10	BE018547	BE018547 bh81h06.y	c 466	18	0.6	533	17	AQ468424	AQ468424
c 394	18	0.6	462	10	BE463060	BE463060 EST354355	467	18	0.6	534	12	BG316633	BG316633
c 395	18	0.6	463	9	AU091433	AU091433 AU091433	468	18	0.6	535	10	BE220391	BE220391
c 396	18	0.6	463	13	BI098513	BI098513 IP1_33_D0	c 469	18	0.6	535	12	BG396112	BG396112
c 397	18	0.6	464	12	BF011937	BF011937 us37h06.y	c 470	18	0.6	535	13	BJ208920	BJ208920
c 398	18	0.6	464	12	BF117399	BF117399 uz36f01.y	471	18	0.6	535	14	BQ810173	BQ810173
c 399	18	0.6	465	10	BE471072	BE471072 WHE0283_C	c 472	18	0.6	536	10	BE443512	BE443512
c 400	18	0.6	466	9	AA488220	AA488220 ad08b02.r	c 473	18	0.6	537	9	AA609344	AA609344
c 401	18	0.6	467	10	AW222105	AW222105 EST298916	c 474	18	0.6	537	9	AA609958	AA609958
c 402	18	0.6	468	14	N42261	N42261 yw85e05.r1	475	18	0.6	537	12	BE750691	BE750691
c 403	18	0.6	468	17	AZ424608	AZ424608 1M0204J15	476	18	0.6	537	14	BQ186092	BQ186092
c 404	18	0.6	470	9	AA797071	AA797071 vs95h12.r	477	18	0.6	538	12	BF266267	BF266267
c 405	18	0.6	471	12	BG317135	BG317135 947025G04	478	18	0.6	540	12	BG410486	BG410486
c 406	18	0.6	475	12	BF564561	BF564561 UI-R-C4-a	479	18	0.6	542	10	AW861661	AW861661
c 407	18	0.6	475	12	BG143201	BG143201 1a94h08.y	c 480	18	0.6	542	12	BG591622	BG591622
c 408	18	0.6	475	13	BI052727	BI052727 RCO-GN027	481	18	0.6	542	17	BH233474	BH233474
c 409	18	0.6	476	10	AV626273	AV626273 AV626273	482	18	0.6	543	9	AI189604	AI189604
c 410	18	0.6	481	9	AA111795	AA111795 SW31CA189	483	18	0.6	543	10	BE494118	BE494118
c 411	18	0.6	483	9	AT340161	AT340161 qo79b05.x	c 484	18	0.6	544	13	BI747628	BI747628
c 412	18	0.6	483	9	AA526293	AA526293 ni95d06.s	485	18	0.6	545	10	AW857994	AW857994
c 413	18	0.6	483	9	AA576035	AA576035 nm57f07.s	486	18	0.6	545	13	BI542714	BI542714
c 414	18	0.6	483	14	BM700236	BM700236 UI-E-DW1-	487	18	0.6	546	9	AA795881	AA795881
c 415	18	0.6	484	9	AA745478	AA745478 ny64a05.s	488	18	0.6	546	9	AI279507	AI279507
c 416	18	0.6	484	9	AA910358	AA910358 ok83f12.s	c 489	18	0.6	546	9	AI296467	AI296467
c 417	18	0.6	484	10	AW722681	AW722681 c4h12mm.r	c 490	18	0.6	546	12	BF607495	BF607495
c 418	18	0.6	484	13	BM611036	BM611036 170006591	c 491	18	0.6	546	12	BG312656	BG312656
c 419	18	0.6	485	9	AA504330	AA504330 aa61e10.s	c 492	18	0.6	547	10	AW623020	AW623020
c 420	18	0.6	487	17	BH217098	BH217098 1006051F0	c 493	18	0.6	547	10	BE586853	BE586853
c 421	18	0.6	489	9	AI188204	AI188204 qd66f12.x	c 494	18	0.6	547	12	BF073278	BF073278
c 422	18	0.6	489	14	W34918	W34918 mc60g10.r1	c 495	18	0.6	548	12	BG887745	BG887745
c 423	18	0.6	490	12	BF436933	BF436933 7p56c02.x	c 496	18	0.6	549	9	AI608132	AI608132
c 424	18	0.6	490	17	BH627981	BH627981 1007078G0	497	18	0.6	551	17	AZ344175	AZ344175
c 425	18	0.6	491	10	BE334969	BE334969 us91c04.y	c 498	18	0.6	552	12	BF586260	BF586260
c 426	18	0.6	492	10	BE518399	BE518399 WHE0815_F	c 499	18	0.6	554	12	BF320458	BF320458
c 427	18	0.6	495	9	AI636631	AI636631 ts92a04.x	500	18	0.6	557	17	TA277H10Q	TA277H10Q
c 428	18	0.6	495	9	AT9898139	AT9898139 EST267582	501	18	0.6	558	10	AW747579	AW747579
c 429	18	0.6	495	9	AT945998	AT945998 bs19d03.y	502	18	0.6	558	12	BG191235	BG191235
c 430	18	0.6	495	10	AW535515	AW535515 UI-R-C4-a	c 503	18	0.6	558	17	BM664995	BM664995
c 431	18	0.6	495	13	BI725114	BI725114 1031077A1	504	18	0.6	558	17	BH116351	BH116351
c 432	18	0.6	496	9	AA823043	AA823043 vw39h11.r	c 505	18	0.6	560	13	BI181668	BI181668
c 433	18	0.6	496	14	BQ168065	BQ168065 WHE0031.E	c 506	18	0.6	560	13	BI184776	BI184776
c 434	18	0.6	497	10	BB713501	BB713501 BB713501	c 507	18	0.6	560	14	N36019	N36019
c 435	18	0.6	499	9	AI529922	AI529922 ui87a12.y	508	18	0.6	561	9	AI186453	AI186453
c 436	18	0.6	499	10	AW722665	AW722665 c4f10mm.x	c 509	18	0.6	561	17	BH769593	BH769593
c 437	18	0.6	499	12	BG396040	BG396040 602458779	c 510	18	0.6	562	10	BE443310	BE443310
c 438	18	0.6	499	17	AZ033668	AZ033668 RPCI-23-3	511	18	0.6	563	9	AJ443060	AJ443060
c 439	18	0.6	501	9	AI897634	AI897634 EST267077	512	18	0.6	564	9	AI194352	AI194352
c 440	18	0.6	502	10	BE481701	BE481701 167173.BA	513	18	0.6	564	17	AQ837843	AQ837843
c 441	18	0.6	504	10	BE587912	BE587912 WHE0665.C	c 514	18	0.6	566	13	BG947217	BG947217
c 442	18	0.6	508	13	BI718806	BI718806 1031031F0	c 515	18	0.6	568	13	BI1996621	BI1996621
c 443	18	0.6	508	14	BM799338	BM799338 K-EST0083	c 516	18	0.6	569	13	BM191050	BM191050
c 444	18	0.6	510	10	AW762147	AW762147 ur54c05.y	517	18	0.6	569	17	AZ617784	AZ617784



518	18	0.6	570	12	BG815845	BG815845 dad74c11.	591	18	0.6	632	13	BI855623	BI855623 603383128
c 519	18	0.6	572	14	AA031174	AA031174 mf148f09.r	592	18	0.6	632	14	BM949826	BM949826 UI-M-EGOp
c 520	18	0.6	572	9	WB3563	WB3563 mi29c10.r1	c 593	18	0.6	634	13	BG922229	BG922229 602820012
c 521	18	0.6	573	10	BE352019	BE352019 894054G03	c 594	18	0.6	634	13	BI221817	BI221817 602937080
c 522	18	0.6	573	12	BF732233	BF732233 nae08a05.	c 595	18	0.6	635	10	AW565147	AW565147 LG1_328_H
c 523	18	0.6	575	10	BE426467	BE426467 WHE0335_D	c 596	18	0.6	636	9	AI734328	AI734328 606029F11
c 524	18	0.6	575	12	BG446435	BG446435 GA_Eb003	c 597	18	0.6	636	9	AL644649	AL644649 AL644649
c 525	18	0.6	575	14	C78684	C78684 C78684 Mous	c 598	18	0.6	636	13	BI718554	BI718554 1031028E0
c 526	18	0.6	576	12	AA021851	AA021851 mh82c10.r	c 599	18	0.6	636	13	BI718554	BI718554 1031028E0
c 527	18	0.6	576	12	BG285981	BG285981 602381158	c 600	18	0.6	638	12	BF164209	BF164209 601773155
c 528	18	0.6	579	17	BI721597	BI721597 1031056H0	c 601	18	0.6	639	10	BM650274	BM650274 BB650274
c 529	18	0.6	579	17	AQ49450	AQ49450 mgx00023J	c 602	18	0.6	639	10	BM661282	BM661282 BB661282
c 530	18	0.6	581	12	BF429279	BF429279 WHE1801_F	c 603	18	0.6	642	13	BM642443	BM642443 170006873
c 531	18	0.6	581	14	W30302	W30302 mc28d03.r1	c 604	18	0.6	645	17	BE407535	BE407535 601300382
c 532	18	0.6	582	9	AA510905	AA510905 vH60b11.r	c 605	18	0.6	645	17	AZ337868	AZ337868 IM068011
c 533	18	0.6	583	9	AI833790	AI833790 605095D11	c 606	18	0.6	645	17	BI3999386	BI3999386 AG-ND-140
c 534	18	0.6	583	14	BM839874	BM839874 K-EST0116	c 607	18	0.6	647	9	AU236697	AU236697 AU236697
c 535	18	0.6	584	12	BF035496	BF035496 601454788	c 608	18	0.6	647	13	BI919702	BI919702 EST359637
c 536	18	0.6	586	13	BI360436	BI360436 387343 MA	c 609	18	0.6	647	17	AQ938910	AQ938910 NLI-DL8R
c 537	18	0.6	588	12	BG086714	BG086714 H3129F11-	c 610	18	0.6	650	17	B26224	B26224 TS8TFB TAM
c 538	18	0.6	588	14	BQ117267	BQ117267 EST602843	c 611	18	0.6	651	12	BG723040	BG723040 602695579
c 539	18	0.6	588	14	BQ554291	BQ554291 H4027C04-	c 612	18	0.6	652	12	BF116020	BF116020 7u76d10.x
c 540	18	0.6	589	12	BG875929	BG875929 IL3-CT022	c 613	18	0.6	652	17	BH030656	BH030656 RPC1-24-2
c 541	18	0.6	590	9	AA092204	AA092204 mh02602.r	c 614	18	0.6	654	10	BE389861	BE389861 601284368
c 542	18	0.6	590	9	AA273515	AA273515 vc38a02.r	c 615	18	0.6	655	17	AZ371972	AZ371972 IM0123E11
c 543	18	0.6	590	12	BE759313	BE759313 an_3667 A	c 616	18	0.6	656	10	BE291771	BE291771 BB291771
c 544	18	0.6	591	12	BG821417	BG821417 602724873	c 617	18	0.6	656	12	BG483701	BG483701 602503556
c 545	18	0.6	592	9	AU144868	AU144868 AU144868	c 618	18	0.6	659	10	AW221934	AW221934 EST298745
c 546	18	0.6	592	12	BF037105	BF037105 601456871	c 619	18	0.6	660	17	CNS02557	AL181636 Tetraodon
c 547	18	0.6	592	12	BF177019	BF177019 EMI_2_E09	c 620	18	0.6	661	13	BG966873	BG966873 602834374
c 548	18	0.6	595	9	AI296538	AI296538 LP10481.5	c 621	18	0.6	662	13	BI193833	BI193833 602948330
c 549	18	0.6	596	14	C78675	C78675 C78675 Mous	c 622	18	0.6	662	13	BM592388	BM592388 170006874
c 550	18	0.6	597	12	BG333464	BG333464 602460113	c 623	18	0.6	663	12	BG453166	BG453166 NF089B04L
c 551	18	0.6	597	12	AW211329	AW211329 uo79f10.y	c 624	18	0.6	664	14	BO258352	BO258352 NISC_Kp12
c 552	18	0.6	598	17	AZ970992	AZ970992 2M0244013	c 625	18	0.6	665	14	BM938713	BM938713 UI-M-CGOp
c 553	18	0.6	600	9	AA275796	AA275796 vc25h11.r	c 626	18	0.6	665	17	AZ392170	AZ392170 IM0154K16
c 554	18	0.6	600	12	BG801359	BG801359 0102-57 M	c 627	18	0.6	667	10	BE602091	BE602091 HVSMCH010
c 555	18	0.6	600	13	BI991681	BI991681 4104-04 M	c 628	18	0.6	668	12	BG703583	BG703583 602686419
c 556	18	0.6	601	13	BI728601	BI728601 1031100C0	c 629	18	0.6	668	13	BJ497081	BJ497081 BJ497081
c 557	18	0.6	605	12	BF702701	BF702701 MI-P-E3-a	c 630	18	0.6	669	12	BG598954	BG598954 EST503854
c 558	18	0.6	606	10	AW961258	AW961258 EST373330	c 631	18	0.6	670	14	BM977765	BM977765 UI-CF-EN1
c 559	18	0.6	606	14	BQ564654	BQ564654 9120f12.y	c 632	18	0.6	670	17	BH004891	BH004891 BMBAC08M1
c 560	18	0.6	608	9	AA239467	AA239467 my20808.r	c 633	18	0.6	670	17	AG127148	AG127148 Pan trogl
c 561	18	0.6	609	13	BI144499	BI144499 602908688	c 634	18	0.6	671	10	BB414804	BB414804 BB414804
c 562	18	0.6	609	17	AZ269297	AZ269297 RPC1-23-7	c 635	18	0.6	672	14	BQ019716	BQ019716 UI-H-ED0-
c 563	18	0.6	609	17	AG161306	AG161306 Pan trogl	c 636	18	0.6	672	14	BQ019716	BQ019716 UI-H-ED0-
c 564	18	0.6	611	13	BM670515	BM670515 UI-E-DM1-	c 637	18	0.6	674	10	BQ8064612	BB554612 BB554612
c 565	18	0.6	612	10	AW741075	AW741075 uq94e06.y	c 638	18	0.6	675	14	BQ806921	BQ806921 WHE3584_F
c 566	18	0.6	615	9	AL680055	AL680055 AL680055	c 639	18	0.6	677	14	BQ411837	BQ411837 CA_Ed004
c 567	18	0.6	617	9	BI519863	BI519863 LD39687.5	c 640	18	0.6	677	14	BQ473104	BQ473104 rc84a11.y
c 568	18	0.6	617	10	BB616288	BB616288 BB616288	c 641	18	0.6	677	14	BQ837962	BQ837962 WHE2905_A
c 569	18	0.6	619	9	AL628932	AL628932 AL628932	c 642	18	0.6	680	17	AZ095325	AZ095325 RPC1-23-1
c 570	18	0.6	620	10	BE392600	BE392600 601307424	c 643	18	0.6	682	13	BM321369	BM321369 rockefell
c 571	18	0.6	621	12	BF153194	BF153194 010P07 Ma	c 644	18	0.6	682	9	AL662455	AL662455 AL662455
c 572	18	0.6	621	13	BQ096108	BQ096108 BJ096108	c 645	18	0.6	683	13	BG923134	BG923134 602823962
c 573	18	0.6	622	12	BG130752	BG130752 EST463644	c 646	18	0.6	683	13	BJ136784	BJ136784 BJ136784
c 574	18	0.6	623	10	BE390292	BE390292 601285181	c 647	18	0.6	683	17	AZ323647	AZ323647 IM0045M10
c 575	18	0.6	623	10	BE435420	BE435420 EST406498	c 648	18	0.6	685	9	AU213990	AU213990 AU213990
c 576	18	0.6	623	10	BE443048	BE443048 WHE1114.B	c 649	18	0.6	685	10	BB042440	BB042440 BB042440
c 577	18	0.6	624	13	BI110309	BI110309 602900740	c 650	18	0.6	685	12	BF117815	BF117815 uz08c08.y
c 578	18	0.6	626	9	AI667826	AI667826 605028H07	c 651	18	0.6	685	14	BQ05974	BQ05974 QCG9U14.y
c 579	18	0.6	626	17	AZ000930	AZ000930 RPC1-23-3	c 652	18	0.6	686	14	BQ789194	BQ789194 WHE158_F
c 580	18	0.6	626	17	AZ000930	AZ000930 RPC1-23-3	c 653	18	0.6	686	14	BQ789194	BQ789194 WHE158_F
c 581	18	0.6	627	10	BB656104	BB656104 BB656104	c 654	18	0.6	687	13	BM448830	BM448830 DSA028G02
c 582	18	0.6	628	9	AI297117	AI297117 LP11337.5	c 655	18	0.6	688	13	W76810	W76810 me83f07.r1
c 583	18	0.6	628	9	AU215688	AU215688 AU215688	c 656	18	0.6	692	13	BI552119	BI552119 603195061
c 584	18	0.6	628	13	BG915757	BG915757 WHE1258_E	c 657	18	0.6	693	10	BE378712	BE378712 601237147
c 585	18	0.6	628	13	BG915757	BG915757 602814288	c 658	18	0.6	694	14	BM840013	BM840013 K-EST0117
c 586	18	0.6	629	10	AW414162	AW414162 uo93f01.y	c 659	18	0.6	697	14	BM680291	BM680291 UI-E-E01-
c 587	18	0.6	629	13	BM411907	BM411907 EST586234	c 660	18	0.6	701	10	AV751438	AV751438 AV751438
c 588	18	0.6	632	10	BM662463	BM662463 BM662463	c 661	18	0.6	701	17	BI151263	BI151263 602917231
c 589	18	0.6	632	12	BG031712	BG031712 602300868	c 662	18	0.6	701	17	BI151263	BI151263 602917231
c 590	18	0.6	632	13	BI725113	BI725113 1031077A1	c 663	18	0.6	702	12	BG764245	BG764245 602736116
										705	13	BI560864	BI560864 603254033

664	18	705	13	BM020234	BM020234	603648828	c 737	18	0.6	820	11	AK020874	Mus musc
665	18	706	13	B1410583	B1410583	602964143	c 738	18	0.6	825	13	B1304659	
666	18	707	13	BM624639	BM624639	170006874	739	18	0.6	835	9	AA829751	od35ei12.s
667	18	708	12	BG862149	BG862149	602795716	c 740	18	0.6	836	10	BE178731	
668	18	709	17	BH084250	BH084250	RP01-24-3	c 741	18	0.6	837	12	BG473084	
669	18	710	12	BF167713	BF167713	601774423	c 742	18	0.6	841	13	BI691860	
670	18	711	13	B1850754	B1850754	imageqcl1	c 743	18	0.6	845	9	AL559252	AL559252
671	18	712	10	BE636604	BE636604	rockefell	c 744	18	0.6	845	12	BF788104	
672	18	713	13	B1823892	B1823892	603039059	745	18	0.6	846	13	BM359463	GA_Ea002
673	18	716	10	AV707602	AV707602	AV707602	746	18	0.6	849	12	BF666644	
674	18	718	12	BF341602	BF341602	602016022	747	18	0.6	852	12	BF314916	
675	18	719	10	AW940604	AW940604	GHI5883.3	c 748	18	0.6	855	14	BQ505020	
676	18	719	12	BF489097	BF489097	AT24801.5	c 749	18	0.6	856	13	B1146833	
677	18	720	10	AW220283	AW220283	EST3029766	750	18	0.6	857	17	AQ747018	HS_5538_A
678	18	721	12	BF867608	BF867608	963092H09	c 751	18	0.6	860	12	BG328120	602427176
679	18	722	12	BG758156	BG758156	6027112325	752	18	0.6	860	13	BI854946	
680	18	724	13	BM407663	BM407663	EST581990	753	18	0.6	862	12	BG723207	602690812
681	18	725	12	BF182127	BF182127	601804667	c 754	18	0.6	867	13	BI655661	603283948
682	18	728	12	BF489125	BF489125	AT24834.5	755	18	0.6	869	12	BF666333	602119251
683	18	728	17	AG033889	AG033889	pan trogl	c 756	18	0.6	871	9	AL526108	
684	18	729	13	BI771203	BI771203	603059213	757	18	0.6	872	13	BM019773	
685	18	731	13	BI733584	BI733584	603352437	c 758	18	0.6	873	9	AL539772	
686	18	733	9	AA177448	AA177448	vu19e06.r	c 759	18	0.6	873	13	BI414418	
687	18	733	14	BO575609	BO575609	UI-H-EZ1-	c 760	18	0.6	874	13	BI655236	
688	18	735	12	BG173068	BG173068	602335552	c 761	18	0.6	876	9	AL537290	
689	18	736	17	CNS0405P	AL307366	Tetraodon	762	18	0.6	876	12	BF217835	
690	18	739	9	AA790490	AA790490	vw05b05.r	763	18	0.6	878	12	BG355119	602324814
691	18	739	17	AO688874	AO688874	nrbxb0078B	c 764	18	0.6	880	12	BF537713	
692	18	740	10	AV764019	AV764019	AV764019	c 765	18	0.6	881	13	BM019895	
693	18	743	12	BG887487	BG887487	EST513338	c 766	18	0.6	882	13	BI853668	
694	18	744	13	BI309716	BI309716	EST531126	c 767	18	0.6	883	13	BI408049	
695	18	745	9	AJ449216	AJ449216		768	18	0.6	883	17	CNS01HSP	
696	18	749					c 769	18	0.6	885	17	CNS04K4R	
697	18	749	12	BG772601	BG772601	602720706	c 770	18	0.6	885	17	CNS04NCT	
698	18	751	14	BQ689942	BQ689942	AGENCOURT	c 771	18	0.6	886	9	AL539785	
699	18	752	11	AK006944	AK006944	Mus muscu	c 771	18	0.6	886	9	AL539785	
700	18	752	12	BE962516	BE962516	601655825	c 772	18	0.6	886	10	BE542860	
701	18	752	13	BI694625	BI694625	603347848	c 773	18	0.6	890	12	BF160408	6010767679
702	18	754	12	BG468654	BG468654	602510258	c 774	18	0.6	892	12	BG714423	
703	18	755	9	AL537385	AL537385		c 775	18	0.6	895	14	BQ963594	AGENCOURT
704	18	756	13	BI733912	BI733912	6033555623	c 776	18	0.6	899	14	BQ924538	AGENCOURT
705	18	756	17	BH674842	BH674842	BOMOF21TF	c 777	18	0.6	903	17	AZ691717	ENTP071TR
706	18	765	12	BE797787	BE797787	601586388	778	18	0.6	905	13	BI175496	OSTR047E2
707	18	767	10	BE377558	BE377558	601229746	c 779	18	0.6	906	14	BQ677208	AGENCOURT
708	18	768	12	BF631334	BF631334	HVSMEB001	c 780	18	0.6	906	14	BQ960087	AGENCOURT
709	18	768	12	BG887486	BG887486	EST513337	c 781	18	0.6	907	14	BQ945752	AGENCOURT
710	18	769	17	BE389523	BE389523	601282874	c 782	18	0.6	907	17	CNS02PPF	Tetraodon
711	18	770	9	AL518386	AL518386	AL518386	c 783	18	0.6	908	17	CNS02CGW	Tetraodon
712	18	772	17	CNS046LD	CNS046LD	Tetraodon	c 784	18	0.6	909	13	BI904890	Tetraodon
713	18	774	10	BE282936	BE282936	601100409	785	18	0.6	909	13	BI904890	603168926
714	18	775	12	BG474100	BG474100	602516671	786	18	0.6	912	12	BG116981	602319048
715	18	776	13	BI652143	BI652143	603299152	787	18	0.6	915	14	BQ648913	AGENCOURT
716	18	777	12	BG243046	BG243046	602353536	c 788	18	0.6	916	14	BQ963508	AGENCOURT
717	18	779	10	BE266129	BE266129	601190947	c 789	18	0.6	923	14	BQ425118	AGENCOURT
718	18	779	12	BF535768	BF535768	602051273	c 790	18	0.6	924	12	BQ900972	AGENCOURT
719	18	781	17	CNS077KV	AL432821	T3 end of	c 791	18	0.6	926	14	BQ686999	AGENCOURT
720	18	785	12	BG066275	BG066275	H3041A09-	c 792	18	0.6	932	12	BE796901	601587255
721	18	785	13	BI415864	BI415864	602988072	c 793	18	0.6	933	17	BH147131	ENTP071TR
722	18	786	13	BM411890	BM411890	EST586217	c 794	18	0.6	941	14	BQ676180	AGENCOURT
723	18	789	13	AG915330	BG915330	602814655	c 795	18	0.6	944	17	CNS02JCB	Tetraodon
724	18	794	9	AA796623	AA796623	vp31q03.r	796	18	0.6	948	12	BF789245	602105144
725	18	794	12	BE972132	BE972132	601651591	c 797	18	0.6	949	12	BF234946	602027390
726	18	794	17	BH330657	BH330657	CH230-125	c 798	18	0.6	949	13	BI825673	603072655
727	18	798	12	BF127237	BF127237	601650894	799	18	0.6	953	12	BG469969	602533441
728	18	799	14	BQ953020	BQ953020	AGENCOURT	c 800	18	0.6	955	14	AL558823	AL558823
729	18	802	11	AY104906	AY104906	zea may	c 801	18	0.6	956	9	AL558823	AGENCOURT
730	18	804	12	BF260468	BF260468	HVSMEF002	c 802	18	0.6	959	12	BG863102	602798160
731	18	807	12	BF204907	BF204907	601866720	c 803	18	0.6	959	14	BQ922851	AGENCOURT
732	18	807	13	BM048518	BM048518	603625032	c 804	18	0.6	964	13	BI552167	603195216
733	18	808	13	BG914937	BG914937	602813795	c 805	18	0.6	964	14	BQ942017	AGENCOURT
734	18	812	13	BI105207	BI105207	6028993483	c 806	18	0.6	966	13	BI656854	603284151
735	18	814	13	BI698032	BI698032	602334182	c 807	18	0.6	969	12	BG247627	602359313
736	18	816	13	BI150497	BI150497	6029315153	c 808	18	0.6	969	12	BG256134	602367263

c 810	18	0.6	969	17	CNS06W16	AL417856 T3 end of	c 883	17	0.5	109	12	BE842535	BE842535 IL3-ST014
c 811	18	0.6	971	14	BQ954296	BQ954296 AGENCOURT	c 884	17	0.5	110	12	BF871664	BF871664 QV0-ET015
c 812	18	0.6	974	9	AL551463	AL551463 AGENCOURT	c 885	17	0.5	111	13	BI727422	BI727422 IL3-10T02C0
c 813	18	0.6	974	17	CNS03E10	AL240417 Tetraodon	c 886	17	0.5	112	12	BE842522	BE842522 IL3-ST014
c 814	18	0.6	977	13	BI4111597	BI4111597 602965824	c 887	17	0.5	112	12	BE842524	BE842524 IL3-ST014
c 815	18	0.6	985	14	BQ877325	BQ877325 AGENCOURT	c 888	17	0.5	113	12	BE842537	BE842537 IL3-ST014
c 816	18	0.6	985	17	CNS01UPO	AL168117 Tetraodon	c 889	17	0.5	113	12	BE842532	BE842532 IL3-ST014
c 817	18	0.6	989	12	BG332198	BG332198 602433528	c 890	17	0.5	114	12	BE842534	BE842534 IL3-ST014
c 818	18	0.6	993	17	CNS00D0T	AL075286 Drosophila	c 891	17	0.5	114	12	BE842534	BE842534 IL3-ST014
c 819	18	0.6	1004	13	BI551552	BI551552 603194394	c 892	17	0.5	115	12	BE988634	BE988634 UI-M-CG0p
c 820	18	0.6	1007	17	CNS04ODF	AL268764 Tetraodon	c 893	17	0.5	117	14	H92039	H92039 ys82e12.s1
c 821	18	0.6	1012	9	AL530671	AL530671 AL530671	c 894	17	0.5	117	14	BQ458388	BQ458388 HA05G04f
c 822	18	0.6	1022	14	BQ964327	BQ964327 AGENCOURT	c 894	17	0.5	126	14	BQ458388	BQ458388 HA05G04f
c 823	18	0.6	1024	17	CNS03OPP	AL253654 Tetraodon	c 895	17	0.5	127	12	BE983793	BE983793 UI-M-CG0p
c 824	18	0.6	1025	13	BM557677	BM557677 AGENCOURT	c 896	17	0.5	130	12	BF460938	BF460938 UI-M-CG0p
c 825	18	0.6	1026	14	BQ709583	BM557677 AGENCOURT	c 897	17	0.5	132	12	BF464665	BF464665 UI-M-CG0p
c 826	18	0.6	1030	17	CNS04CBW	BQ709583 AGENCOURT	c 898	17	0.5	132	12	BF465949	BF465949 UI-M-CG0p
c 827	18	0.6	1040	17	CNS03Q03	AL284261 Tetraodon	c 899	17	0.5	133	9	AA731348	AA731348 n294d10.s
c 828	18	0.6	1045	13	BI658550	AL267924 Tetraodon	c 900	17	0.5	133	9	AA827636	AA827636 OD55902.s
c 829	18	0.6	1047	13	BM544499	BI658550 603283777	c 901	17	0.5	136	9	AA139263	AA139263 mq02e06.f
c 830	18	0.6	1049	12	BF577950	BM544499 AGENCOURT	c 902	17	0.5	136	9	AA173307	AA173307 AV173307
c 831	18	0.6	1050	9	AV214969	BF577950 602091972	c 903	17	0.5	137	9	AA724434	AA724434 ab91g11.s
c 832	18	0.6	1052	17	CNS00ITV	AV214969 AV214969	c 904	17	0.5	140	10	AW197526	AW197526 xm39f06.x
c 833	18	0.6	1065	13	BI411015	AL075434 Drosophila	c 905	17	0.5	140	12	BG012637	BG012637 QV3-GN033
c 834	18	0.6	1068	14	BM909849	BI411015 602963370	c 906	17	0.5	141	17	BH223560	BH223560 1006113F0
c 835	18	0.6	1071	13	BM560489	BM909849 AGENCOURT	c 907	17	0.5	141	9	AA743332	AA743332 ny22f03.s
c 836	18	0.6	1075	14	BM925478	BM560489 AGENCOURT	c 908	17	0.5	147	14	BQ636167	BQ636167 hd05d05.y
c 837	18	0.6	1085	12	BF578325	BM925478 AGENCOURT	c 909	17	0.5	148	12	BF173209	BF173209 MYE0651a
c 838	18	0.6	1088	12	BF209692	BF578325 602092683	c 910	17	0.5	149	12	BF817967	BF817967 MR2-C1012
c 839	18	0.6	1091	14	BQ944885	BF209692 601874226	c 911	17	0.5	151	17	BH199133	BH199133 TC3-70M11
c 840	18	0.6	1092	17	CNS05MXA	BQ944885 AGENCOURT	c 912	17	0.5	152	9	AA231901	AA231901 RZ543.R.c
c 841	18	0.6	1093	17	CNS03KIQ	AL344647 Tetraodon	c 913	17	0.5	156	9	AV033363	AV033363 AV033363
c 842	18	0.6	1097	13	BG913347	AL344647 Tetraodon	c 914	17	0.5	157	12	BE842527	BE842527 IL3-ST014
c 843	18	0.6	1098	11	AX111914	BG913347 602812071	c 915	17	0.5	157	12	BE842527	BE842527 IL3-ST014
c 844	18	0.6	1098	12	BG248147	AX111914 Zea mays	c 916	17	0.5	158	9	AA805512	AA805512 OG12f04.s
c 845	18	0.6	1101	17	CNS008Q0	BG248147 602400569	c 917	17	0.5	162	9	AL842063	AL842063 AL842063
c 846	18	0.6	1101	17	CNS008Q0	AL052289 Drosophila	c 918	17	0.5	163	9	AI842244	AI842244 UI-M-A11
c 847	18	0.6	1124	14	BQ923373	AL068709 Drosophila	c 919	17	0.5	163	9	AI842244	AI842244 UI-M-A11
c 848	18	0.6	1129	14	BM911004	BQ923373 AGENCOURT	c 920	17	0.5	165	10	AW603908	AW603908 RC2-CN005
c 849	18	0.6	1130	14	BQ684644	BM911004 AGENCOURT	c 921	17	0.5	166	10	AW854556	AW854556 RC3-CT025
c 850	18	0.6	1131	14	BQ276399	BQ684644 AGENCOURT	c 922	17	0.5	166	13	BI508503	BI508503 603059269
c 851	18	0.6	1131	17	CNS03L5Y	BQ276399 AGENCOURT	c 923	17	0.5	166	17	BH224557	BH224557 1006119G0
c 852	18	0.6	1147	14	BQ071836	AL249055 Tetraodon	c 924	17	0.5	169	9	AA928272	AA928272 on79c06.s
c 853	18	0.6	1150	14	BQ926132	AL249055 Tetraodon	c 925	17	0.5	170	9	AA014090	AA014090 mh28d03.f
c 854	18	0.6	1179	12	BF181455	BQ071836 AGENCOURT	c 926	17	0.5	178	13	BG956065	BG956065 CM4-CT065
c 855	18	0.6	1186	17	AG032524	BF181455 601805645	c 927	17	0.5	178	17	TA170A12P	TA170A12P T. brucei
c 856	18	0.6	1190	14	BM922179	AG032524 Pan trogl	c 928	17	0.5	180	9	AJ467324	AJ467324 AJ467324
c 857	18	0.6	1196	13	BM546078	BM922179 AGENCOURT	c 929	17	0.5	180	9	AJ468324	AJ468324 AJ468324
c 858	18	0.6	1212	13	BM925803	BM546078 AGENCOURT	c 930	17	0.5	180	17	BH196416	BH196416 TC3-34F10
c 859	18	0.6	1328	11	AK011107	BM925803 AGENCOURT	c 931	17	0.5	181	12	BE842526	BE842526 IL3-ST014
c 860	18	0.6	1337	12	BG471056	AK011107 Mus muscu	c 932	17	0.5	184	14	D22017	D22017 RIC10187A
c 861	18	0.6	1338	14	BQ684684	BG471056 602512043	c 933	17	0.5	185	10	BE492632	BE492632 WHE0559.C
c 862	18	0.6	1359	14	BM805497	BQ684684 AGENCOURT	c 934	17	0.5	186	9	AA767847	AA767847 OG63f05.S
c 863	18	0.6	1417	14	BM914673	BM805497 AGENCOURT	c 935	17	0.5	186	9	AV094674	AV094674 AV094674
c 864	18	0.6	1434	11	AY106315	BM914673 AGENCOURT	c 936	17	0.5	186	17	AZ824339	AZ824339 2M0098E20
c 865	18	0.6	1444	14	BM921764	AY106315 Zea mays	c 937	17	0.5	186	17	AG0305870	AG0305870 fm13c05.x
c 866	18	0.6	1486	17	AG182550	BM921764 AGENCOURT	c 938	17	0.5	187	12	BA040800	BA040800 zf07g02.s
c 867	18	0.6	1488	13	BM556903	AG182550 Pan trogl	c 939	17	0.5	188	9	AA785083	AA785083 q4g11a1.f
c 868	18	0.6	1489	13	BM467281	BM556903 AGENCOURT	c 940	17	0.5	190	9	AU077190	AU077190 AU077190
c 869	18	0.6	1547	14	BM911629	BM467281 AGENCOURT	c 941	17	0.5	191	10	AW862068	AW862068 RC3-CT034
c 870	18	0.6	1568	13	BM048762	BM911629 AGENCOURT	c 942	17	0.5	192	10	BE842523	BE842523 IL3-ST014
c 871	18	0.6	1780	12	BF581621	BM048762 603625393	c 943	17	0.5	193	12	BE842523	BE842523 IL3-ST014
c 872	18	0.6	2420	10	BE402959	BF581621 602099973	c 944	17	0.5	195	14	BQ694464	BQ694464 1000607.H
c 873	18	0.6	2898	11	AK004737	BE402959 GBX002.G1	c 945	17	0.5	198	9	AA888766	AA888766 am032h12.s
c 874	17	0.5	54	12	BE783642	AK004737 Mus muscu	c 946	17	0.5	198	9	AV172966	AV172966 AV172966
c 875	17	0.5	80	9	AA761381	BE783642 601471219	c 947	17	0.5	198	10	BE269372	BE269372 BB269372
c 876	17	0.5	84	9	AA848081	AA761381 n225q12.s	c 948	17	0.5	201	13	BI127995	BI127995 G069P46Y
c 877	17	0.5	91	9	AA889872	AA848081 OD55b11.s	c 949	17	0.5	203	9	AA161523	AA161523 MSAFCE9G0
c 878	17	0.5	100	9	AA208122	AA889872 al52c06.s	c 950	17	0.5	207	9	AI236315	AI236315 EST232877
c 879	17	0.5	101	12	BE842533	AA208122 mv93d11.f	c 951	17	0.5	208	14	DA2622	DA2622 D42622
c 880	17	0.5	101	9	AA765673	BE842533 IL3-ST014	c 952	17	0.5	210	9	AA767490	AA767490 oa90a10.s
c 881	17	0.5	105	9	AA504849	AA765673 oa04f12.s	c 953	17	0.5	210	9	AV294428	AV294428 AV294428
c 882	17	0.5	109	12	BE842525	AA504849 aa64f06.s	c 954	17	0.5	210	10	AW114475	AW114475 rs51q07.y
						BE842525 IL3-ST014	c 955	17	0.5	210	10	BB291607	BB291607 BB291607



```

source
1. .365
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="Drosophila melanogaster P(SUPor-P) P element
insertion lines"
/note="Inverse PCR was performed on Drosophila
melanogaster strains each of which contains one or more
P(SUPor-P) P-element transposon insertion. The resultant
fragment for each strain was directly sequenced to
determine the genomic sequence at the site of insertion.
Details of the protocols used can be found at
http://www.fruitfly.org/about/methods/inverse.pcr.html."
BASE COUNT      62 a      101 g      118 t
ORIGIN
Query Match      0.8%; Score 25; DB 17; Length 365;
Best Local Similarity 100.0%; Pred. No. 0.41;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 143 GCAGCAACAGCAGCAACAAACAGC 167
      ||||||||||||||||||||||||
Db 153 GCAGCAACAGCAGCAACAAACAGC 129

RESULT 3
BI241907
LOCUS
DEFINITION
RE39108.5prime RE Drosophila melanogaster normalized Embryo pFlc-1
Drosophila melanogaster cDNA clone RE39108 5 similar to CG3036;
FBan003036 'transporter' located on: 2L 25B5-25B6; 05/12/2001,
mRNA sequence.
ACCESSION      BI241907
VERSION
KEYWORDS
SOURCE
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 431)
AUTHORS
Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S.,
Mundall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
Phouanavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and Rubin
,G.M.
BDGP/HMI RE Drosophila EST Project
Unpublished (2001)
Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd
Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
hit genomic AB003575: arm:2L [4533649,4860857]
estimated-cyto:25A2-25B8: 05/12/2001
Plate: RE 391 row: A column: 8
High quality sequence stop: 430.
Location/Qualifiers
1. .431
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="RE Drosophila melanogaster normalized Embryo
pFlc-1"
/sex="male and female"
/dev_stage="0-24 hours mixed stage embryonic"
/lab_host="DH5-alpha Tona"
/note="organ: embryo; Vector: pFlc1; Site:1: XhoI; Site:2:
BamHI; Library was kindly generated by Piero Carninci at
the RIKEN. The library was normalized and excised using
Cre recombinase. Plasmid cDNA library."
BASE COUNT      125 a      100 c      87 g      119 t

FEATURES
source
1. .431
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="RE Drosophila melanogaster normalized Embryo
pFlc-1"
/sex="male and female"
/dev_stage="0-24 hours mixed stage embryonic"
/lab_host="DH5-alpha Tona"
/note="organ: embryo; Vector: pFlc1; Site:1: XhoI; Site:2:
BamHI; Library was kindly generated by Piero Carninci at
the RIKEN. The library was normalized and excised using
Cre recombinase. Plasmid cDNA library."
BASE COUNT      125 a      100 c      87 g      119 t

ORIGIN
Query Match      0.8%; Score 25; DB 13; Length 431;
Best Local Similarity 100.0%; Pred. No. 0.41;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 143 GCAGCAACAGCAGCAACAAACAGC 167
      ||||||||||||||||||||||||
Db 32 GCAGCAACAGCAGCAACAAACAGC 56

RESULT 4
BI482365
LOCUS
DEFINITION
RE65049.5prime RE Drosophila melanogaster normalized Embryo pFlc-1
Drosophila melanogaster cDNA clone RE65049 5 similar to CG3036;
FBan003036 GO:{transporter (GO:0005215)} located on: 2L 25B5-25B6
; 05/16/2001, mRNA sequence.
ACCESSION      BI482365
VERSION
KEYWORDS
SOURCE
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 440)
AUTHORS
Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S.,
Mundall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
Phouanavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and Rubin
,G.M.
BDGP/HMI RE Drosophila EST Project
Unpublished (2001)
Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd
Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
hit genomic AB003575: arm:2L [4533649,4860857]
estimated-cyto:25A2-25B8: 05/16/2001
Plate: RE 650 row: E column: 1
High quality sequence stop: 438.
Location/Qualifiers
1. .440
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="RE Drosophila melanogaster normalized Embryo
pFlc-1"
/sex="male and female"
/dev_stage="0-24 hours mixed stage embryonic"
/lab_host="DH5-alpha Tona"
/note="Organ: embryo; Vector: pFlc1; Site:1: XhoI; Site:2:
BamHI; Library was kindly generated by Piero Carninci at
the RIKEN. The library was normalized and excised using
Cre recombinase. Plasmid cDNA library."
BASE COUNT      130 a      99 c      88 g      123 t

FEATURES
source
1. .440
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="RE Drosophila melanogaster normalized Embryo
pFlc-1"
/sex="male and female"
/dev_stage="0-24 hours mixed stage embryonic"
/lab_host="DH5-alpha Tona"
/note="Organ: embryo; Vector: pFlc1; Site:1: XhoI; Site:2:
BamHI; Library was kindly generated by Piero Carninci at
the RIKEN. The library was normalized and excised using
Cre recombinase. Plasmid cDNA library."
BASE COUNT      130 a      99 c      88 g      123 t

ORIGIN
Query Match      0.8%; Score 25; DB 13; Length 440;
Best Local Similarity 100.0%; Pred. No. 0.41;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 143 GCAGCAACAGCAGCAACAAACAGC 167
      ||||||||||||||||||||||||
Db 31 GCAGCAACAGCAGCAACAAACAGC 55

RESULT 5
AI063687

```

LOCUS AT063687 443 bp mRNA linear EST 19-APR-2001  
DEFINITION GH03640.5prime GH Drosophila melanogaster head pOT2 Drosophila  
melanogaster cDNA clone GH03640 5prime, mRNA sequence.  
ACCESSION A1063687  
VERSION A1063687.1 GI:3339631  
KEYWORDS EST.  
SOURCE fruit fly.  
ORGANISM Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
REFERENCE 1 (bases 1 to 443)  
AUTHORS Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,  
Lewis,S. and Rubin,G.M.  
TITLE BDGP/HMI Drosophila EST Project  
JOURNAL Unpublished (2001)  
COMMENT Contact: Stapleton, M.  
BDGP  
Lawrence Berkeley National Lab  
One Cyclotron Rd, Berkeley, CA 94720, USA  
Fax: 510 486 6798  
Email: <http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu>  
Plate: 36 row: D column: 4  
High quality sequence stop: 344.  
Location/Qualifiers  
1..443  
/organism="Drosophila melanogaster"  
/db\_xref="taxon:7227"  
/clone="GH03640"  
/clone\_lib="GH Drosophila melanogaster head pOT2"  
/sex="male and female"  
/dev\_stage="adult"  
/lab\_host="DH5 - alpha"  
/note="Organ: head; Vector: pOT2; Site\_1: EcoRI; Site\_2:  
XhoI; Sized fractionated cDNAs were directly ligated into  
pOT2. Plasmid cDNA library."  
BASE COUNT 137 a 100 c 85 g 121 t  
ORIGIN  
Query Match 0.8%; Score 25; DB 9; Length 443;  
Best Local Similarity 100.0%; Pred. No. 0.41; Indels 0; Gaps 0;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 143 GCAGCAACGACGAGCAACACAGC 167  
|||||  
DB 11 GCAGCAACGACGACCAACACAGC 35  
|||||  
RESULT 6  
BI365464  
LOCUS  
DEFINITION BI365464 485 bp mRNA linear EST 01-AUG-2001  
Drosophila melanogaster cDNA clone RE50784 5 similar to CG3036;  
FBAN0003036 'transporter' located on: 2L 25B5-25B6;; 05/14/2001,  
mRNA sequence.  
ACCESSION BI365464  
VERSION BI365464.1 GI:15061492  
KEYWORDS EST.  
SOURCE fruit fly.  
ORGANISM Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
REFERENCE 1 (bases 1 to 485)  
AUTHORS Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson  
,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George  
,R., Gonzales,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S.,  
Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,  
Phouanavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and Rubin  
,G.M.  
TITLE BDGP/HMI RE Drosophila EST Project  
JOURNAL Unpublished (2001)  
COMMENT Contact: Stapleton, M.  
BDGP  
Lawrence Berkeley National Lab  
One Cyclotron Rd, Berkeley, CA 94720, USA  
Fax: 510 486 6798  
Email: <http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu>  
hit genomic AE003575: arm:2L [4533649,4860857]  
estimated-cyto:25A2-25B8: 04/11/2001  
Plate: RE.51 row: B column: 3  
High quality sequence stop: 425.  
Location/Qualifiers  
1..491  
/organism="Drosophila melanogaster"  
/db\_xref="taxon:7227"

BDGP  
Lawrence Berkeley National Lab  
One Cyclotron Rd, Berkeley, CA 94720, USA  
Fax: 510 486 6798  
Email: <http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu>  
hit genomic AE003575: arm:2L [4533649,4860857]  
estimated-cyto:25A2-25B8: 05/14/2001  
Plate: RE.507 row: G column: 12  
High quality sequence stop: 420.  
Location/Qualifiers  
1..485  
/organism="Drosophila melanogaster"  
/db\_xref="taxon:7227"  
/clone="RE50784"  
/clone\_lib="RE Drosophila melanogaster normalized Embryo  
pFlc-1"  
/sex="male and female"  
/dev\_stage="0-24 hours mixed stage embryonic"  
/lab\_host="DH5-alpha Tona"  
/note="Organ: embryo; Vector: pFlc1; Site\_1: XhoI; Site\_2:  
BamHI; Library was kindly generated by Piero Carninci at  
the RIKEN. The library was normalized and excised using  
Cre recombinase. Plasmid cDNA library."  
BASE COUNT 143 a 109 c 97 g 136 t  
ORIGIN  
Query Match 0.8%; Score 25; DB 13; Length 485;  
Best Local Similarity 100.0%; Pred. No. 0.41; Indels 0; Gaps 0;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 143 GCAGCAACGACGAGCAACACAGC 167  
|||||  
DB 32 GCAGCAACGACGAGCAACACAGC 56  
|||||  
RESULT 7  
BI165260  
LOCUS  
DEFINITION BI165260 491 bp mRNA linear EST 09-JUL-2001  
Drosophila melanogaster cDNA clone RE05115 5 similar to CG3036;  
FBAN0003036 'transporter' located on: 2L 25B5-25B6;; 04/11/2001,  
mRNA sequence.  
ACCESSION BI165260  
VERSION BI165260.1 GI:14631066  
KEYWORDS EST.  
SOURCE fruit fly.  
ORGANISM Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
REFERENCE 1 (bases 1 to 491)  
AUTHORS Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson  
,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George  
,R., Gonzales,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S.,  
Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,  
Phouanavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and Rubin  
,G.M.  
TITLE BDGP/HMI RE Drosophila EST Project  
JOURNAL Unpublished (2001)  
COMMENT Contact: Stapleton, M.  
BDGP  
Lawrence Berkeley National Lab  
One Cyclotron Rd, Berkeley, CA 94720, USA  
Fax: 510 486 6798  
Email: <http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu>  
hit genomic AE003575: arm:2L [4533649,4860857]  
estimated-cyto:25A2-25B8: 04/11/2001  
Plate: RE.51 row: B column: 3  
High quality sequence stop: 425.  
Location/Qualifiers  
1..491  
/organism="Drosophila melanogaster"  
/db\_xref="taxon:7227"

```

/clone="RE05115"
/db_xref="RE Drosophila melanogaster normalized Embryo
pflc-1"
/sex="male and female"
/dev_stage="0-24 hours mixed stage embryonic"
/lab_host="DH5-alpha Tona"
/notes="Organ: embryo; Vector: pFlc1; Site_1: XhoI; Site_2:
BamHI; Library was kindly generated by Piero Carninci at
the RIKEN. The library was normalized and excised using
Cre recombinase. Plasmid cDNA library."
BASE COUNT      146 a 110 c 96 g 139 t
ORIGIN
Query Match      0.8%; Score 25; DB 13; Length 491;
Best Local Similarity 100.0%; Pred. No. 0.41;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 143 GCAGCAACAGCAGAGCAACACAGC 167
|||||
Db 32 GCAGCAACAGCAGAGCAACACAGC 56

RESULT 8
BI227520
LOCUS
DEFINITION
RE24591.5prime RE Drosophila melanogaster normalized Embryo pFlc-1
Drosophila melanogaster cDNA clone RE24591 5 similar to CG3036;
FBan0003036 'transporter' located on: 2L 25B5-25B6; 04/12/2001,
mRNA sequence.
ACCESSION      BI227520
VERSION        BI227520.1 GI:14694784
KEYWORDS       EST.
SOURCE         fruit fly.
ORGANISM       Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
AUTHORS        Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S.,
Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
Phouanavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and Rubin
,G.M.
TITLE          BDGP/HMI RE Drosophila EST Project
JOURNAL        Unpublished (2001)
COMMENT        Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
hit genomic AB003575; arm:2L [4533649,4860857]
estimated-cyto:25A2-25B8: 04/12/2001
Plate: RE.245 row: H column: 7
High quality sequence stop: 445.
Location/Qualifiers
1. -521
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="RE24591"
/clone_lib="RE Drosophila melanogaster normalized Embryo
pflc-1"
/sex="male and female"
/dev_stage="0-24 hours mixed stage embryonic"
/lab_host="DH5-alpha Tona"
/notes="Organ: embryo; Vector: pFlc1; Site_1: XhoI; Site_2:
BamHI; Library was kindly generated by Piero Carninci at
the RIKEN. The library was normalized and excised using
Cre recombinase. Plasmid cDNA library."
BASE COUNT      154 a 119 c 104 g 144 t
ORIGIN
source
Query Match      0.8%; Score 25; DB 13; Length 526;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 143 GCAGCAACAGCAGAGCAACACAGC 167
|||||
Db 32 GCAGCAACAGCAGAGCAACACAGC 56

RESULT 10
BI356933
LOCUS
DEFINITION
RE43234.5prime RE Drosophila melanogaster normalized Embryo pFlc-1

```

```

Query Match      0.8%; Score 25; DB 13; Length 521;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 143 GCAGCAACAGCAGAGCAACACAGC 167
|||||
Db 32 GCAGCAACAGCAGAGCAACACAGC 56

RESULT 9
BI363907
LOCUS
DEFINITION
RE48884.5prime RE Drosophila melanogaster normalized Embryo pFlc-1
Drosophila melanogaster cDNA clone RE48884 5 similar to CG3036;
FBan0003036 'transporter' located on: 2L 25B5-25B6; 05/13/2001,
mRNA sequence.
ACCESSION      BI363907
VERSION        BI363907.1 GI:15059935
KEYWORDS       EST.
SOURCE         fruit fly.
ORGANISM       Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
AUTHORS        Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S.,
Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
Phouanavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and Rubin
,G.M.
TITLE          BDGP/HMI RE Drosophila EST Project
JOURNAL        Unpublished (2001)
COMMENT        Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
hit genomic AE003575; arm:2L [4533649,4860857]
estimated-cyto:25A2-25B8: 05/13/2001
Plate: RE.488 row: G column: 12
High quality sequence stop: 425.
Location/Qualifiers
1. -526
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="RE48884"
/clone_lib="RE Drosophila melanogaster normalized Embryo
pflc-1"
/sex="male and female"
/dev_stage="0-24 hours mixed stage embryonic"
/lab_host="DH5-alpha Tona"
/notes="Organ: embryo; Vector: pFlc1; Site_1: XhoI; Site_2:
BamHI; Library was kindly generated by Piero Carninci at
the RIKEN. The library was normalized and excised using
Cre recombinase. Plasmid cDNA library."
BASE COUNT      157 a 117 c 108 g 144 t
ORIGIN
source
Query Match      0.8%; Score 25; DB 13; Length 526;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 143 GCAGCAACAGCAGAGCAACACAGC 167
|||||
Db 32 GCAGCAACAGCAGAGCAACACAGC 56

RESULT 10
BI356933
LOCUS
DEFINITION
RE43234.5prime RE Drosophila melanogaster normalized Embryo pFlc-1

```

Drosophila melanogaster cDNA clone RE43234 5 similar to CG3036:  
FBan0003036 'transporter' located on: 2L 25B5-25B6;: 05/13/2001,  
mRNA sequence.

ACCESSION BI356933  
VERSION BI356933.1 GI:15051387  
KEYWORDS EST.  
SOURCE fruit fly.

ORGANISM Drosophila melanogaster  
REFERENCE 1 (bases 1 to 531)  
AUTHORS Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson  
J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George  
R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S.,  
Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,  
Phouanavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and Rubin  
G.M.

TITLE BDGP/HIMI RE Drosophila EST Project

JOURNAL Unpublished (2001)

COMMENT Contact: Stapleton, M.

BDGP Lawrence Berkeley National Lab

One Cyclotron Rd,

Fax: 510 486 6798

Email: [http://www.fruitfly.org/EST\\_est@fruitfly.berkeley.edu](http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu)

hit genomic AE003575: arm:2L [4533649,4860857]

estimated-cyto:25A2-25B8: 05/13/2001

Plate: RE.432 row: C column: 10

High quality sequence stop: 486.

Location/Qualifiers

1..531

/organism="Drosophila melanogaster"

/db\_xref="taxon:7227"

/clone="RE43234"

/clone\_lib="RE Drosophila

pFlc-1"

/sex="male and female"

/dev\_stage="0-24 hours mixed stage embryonic"

/lab\_host="DH5-alpha Tona"

/note="Organ: embryo; Vector: pFlc1; Site\_1: XhoI; Site\_2:  
BamHI; Library was kindly generated by Piero Carninci at  
the RIKEN. The library was normalized and excised using  
Cre recombinase. Plasmid cDNA library."

158 a 119 c 107 g 147 t

BASE COUNT

ORIGIN

Query Match 0.8%; Score 25; DB 13; Length 531;  
Best Local Similarity 100.0%; Pred.No. 0.4;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 143 GCAGCAACAGCAGCAGCAACACAGC 167

|||||

Db 35 GCAGCAACAGCAGCAGCAACACAGC 59

RESULT 11

At114180

LOCUS

DEFINITION GH10888.5prime GH Drosophila melanogaster head pOT2 Drosophila  
melanogaster cDNA clone GH10888 5prime, mRNA sequence.

ACCESSION At114180

VERSION At114180.1 GI:3514983

KEYWORDS EST.

SOURCE fruit fly.

ORGANISM Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 544)

REFERENCE Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,

LEWIS,S. and Rubin,G.M.

TITLE BDGP/HIMI Drosophila EST Project

JOURNAL

COMMENT

Unpublished (2001)

Contact: Stapleton, M.

BDGP

Lawrence Berkeley National Lab

One Cyclotron Rd, Berkeley, CA 94720, USA

Fax: 510 486 6798

Email: [http://www.fruitfly.org/EST\\_est@fruitfly.berkeley.edu](http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu)

Plate: 108 row: H column: 4

High quality sequence stop: 447.

Location/Qualifiers

1..544

/organism="Drosophila melanogaster"

/db\_xref="taxon:7227"

/clone="GH10888"

/clone\_lib="GH Drosophila melanogaster head pOT2"

/sex="male and female"

/dev\_stage="adult"

/lab\_host="DH5 - alpha"

/note="Organ: head; Vector: pOT2; Site\_1: EcoRI; Site\_2:  
XhoI; Sized fractionated cDNAs were directly ligated into  
pOT2. Plasmid cDNA library."

163 a 120 c 116 g 145 t

BASE COUNT

ORIGIN

Query Match 0.8%; Score 25; DB 9; Length 544;

Best Local Similarity 100.0%; Pred.No. 0.4;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 143 GCAGCAACAGCAGCAGCAACACAGC 167

|||||

Db 12 GCAGCAACAGCAGCAGCAACACAGC 36

RESULT 12

At135320

LOCUS

DEFINITION

GH13005.5prime GH Drosophila melanogaster head pOT2 Drosophila

melanogaster cDNA clone GH13005 5prime, mRNA sequence.

ACCESSION At135320

VERSION At135320.1 GI:3627878

KEYWORDS EST.

SOURCE fruit fly.

ORGANISM Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 545)

REFERENCE Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,

LEWIS,S. and Rubin,G.M.

TITLE BDGP/HIMI Drosophila EST Project

JOURNAL

COMMENT

Unpublished (2001)

Contact: Stapleton, M.

BDGP

Lawrence Berkeley National Lab

One Cyclotron Rd, Berkeley, CA 94720, USA

Fax: 510 486 6798

Email: [http://www.fruitfly.org/EST\\_est@fruitfly.berkeley.edu](http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu)

Plate: 130 row: A column: 5

High quality sequence stop: 544.

Location/Qualifiers

1..545

/organism="Drosophila melanogaster"

/db\_xref="taxon:7227"

/clone="GH13005"

/clone\_lib="GH Drosophila melanogaster head pOT2"

/sex="male and female"

/dev\_stage="adult"

/lab\_host="DH5 - alpha"

/note="Organ: head; Vector: pOT2; Site\_1: EcoRI; Site\_2:  
XhoI; Sized fractionated cDNAs were directly ligated into  
pOT2. Plasmid cDNA library."

163 a 121 c 114 g 147 t

BASE COUNT

ORIGIN



```

Query Match      0.8%; Score 25; DB 9; Length 545;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 143 GCAGCAACAGCAGCAGCAACACAGC 167
      |||||||
Db 15 GCAGCAACAGCAGCAGCAACACAGC 39

RESULT 13
BI362358
LOCUS
DEFINITION
  BI362358 560 bp mRNA linear EST 01-AUG-2001
  RE46895.5prime RE Drosophila melanogaster normalized Embryo pFlc-1
  Drosophila melanogaster cDNA clone RE46895.5 similar to CG3036;
  FBan003036 'transporter' located on: 2L 25B5-25B6; 05/13/2001,
  mRNA sequence.
ACCESSION
  BI362358
VERSION
  BI362358.1 GI:15058386
KEYWORDS
  EST.
SOURCE
  fruit fly.
ORGANISM
  Drosophila melanogaster
  Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
  Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
  Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
  1 (bases 1 to 560)
AUTHORS
  Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
  ,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
  ,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S.,
  Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
  Phouanavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and Rubin
  ,G.M.
  BDGP/HMI RE Drosophila EST Project
  Unpublished (2001)
  Contact: Stapleton, M.
  BDGP
  Lawrence Berkeley National Lab
  One Cyclotron Rd, Berkeley, CA 94720, USA
  Fax: 510 486 6798
  Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
  hit genomic AE003575; arm:2L [4533649,4860857]
  estimated-cyto:25A2-25B8; 05/13/2001
  Plate: RE.468 row: H column: 11
  High quality sequence stop: 497.
  Location/Qualifiers
  1..560
  /organism="Drosophila melanogaster"
  /db_xref="taxon:7227"
  /clone_lib="RE Drosophila melanogaster normalized Embryo
  pFlc-1"
  /sex="male and female"
  /dev_stage="0-24 hours mixed stage embryonic"
  /lab_host="DH5-alpha Tona"
  /note="Organ: embryo; Vector: pFlc1; Site_1: XhoI; Site_2:
  BamHI; Library was kindly generated by Piero Carninci at
  the RIKEN. The library was normalized and excised using
  Cre recombinase. Plasmid cDNA library."
  Cre recombinase. Plasmid cDNA library."

BASE COUNT 167 a 123 c 117 g 153 t
ORIGIN
  Query Match      0.8%; Score 25; DB 13; Length 560;
  Best Local Similarity 100.0%; Pred. No. 0.4;
  Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 143 GCAGCAACAGCAGCAGCAACACAGC 167
      |||||||
Db 31 GCAGCAACAGCAGCAGCAACACAGC 55

RESULT 14
BI164055
LOCUS
DEFINITION
  BI164055 577 bp mRNA linear EST 09-JUL-2001
  RE05402.5prime RE Drosophila melanogaster normalized Embryo pFlc-1
  Drosophila melanogaster cDNA clone RE05402.5 similar to CG3036;
  FBan003036 'transporter' located on: 2L 25B5-25B6; 04/11/2001,
  mRNA sequence.
ACCESSION
  BI164055
VERSION
  BI164055.1 GI:14631292
KEYWORDS
  EST.
SOURCE
  fruit fly.
ORGANISM
  Drosophila melanogaster
  Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
  Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
  Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
  1 (bases 1 to 579)
  
```

```

DEFINITION
  RE03625.5prime RE Drosophila melanogaster normalized Embryo pFlc-1
  Drosophila melanogaster cDNA clone RE03625.5 similar to CG3036;
  FBan003036 'transporter' located on: 2L 25B5-25B6; 04/11/2001,
  mRNA sequence.
ACCESSION
  BI164055
VERSION
  BI164055.1 GI:14629861
KEYWORDS
  EST.
SOURCE
  fruit fly.
ORGANISM
  Drosophila melanogaster
  Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
  Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
  Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
  1 (bases 1 to 577)
AUTHORS
  Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
  ,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
  ,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S.,
  Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
  Phouanavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and Rubin
  ,G.M.
  BDGP/HMI RE Drosophila EST Project
  Unpublished (2001)
  Contact: Stapleton, M.
  BDGP
  Lawrence Berkeley National Lab
  One Cyclotron Rd, Berkeley, CA 94720, USA
  Fax: 510 486 6798
  Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
  hit genomic AE003575; arm:2L [4533649,4860857]
  estimated-cyto:25A2-25B8; 04/11/2001
  Plate: RE.36 row: C column: 1
  High quality sequence stop: 497.
  Location/Qualifiers
  1..577
  /organism="Drosophila melanogaster"
  /db_xref="taxon:7227"
  /clone_lib="RE Drosophila melanogaster normalized Embryo
  pFlc-1"
  /sex="male and female"
  /dev_stage="0-24 hours mixed stage embryonic"
  /lab_host="DH5-alpha Tona"
  /note="Organ: embryo; Vector: pFlc1; Site_1: XhoI; Site_2:
  BamHI; Library was kindly generated by Piero Carninci at
  the RIKEN. The library was normalized and excised using
  Cre recombinase. Plasmid cDNA library."
  Cre recombinase. Plasmid cDNA library."

BASE COUNT 168 a 131 c 120 g 158 t
ORIGIN
  Query Match      0.8%; Score 25; DB 13; Length 577;
  Best Local Similarity 100.0%; Pred. No. 0.4;
  Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 143 GCAGCAACAGCAGCAGCAACACAGC 167
      |||||||
Db 32 GCAGCAACAGCAGCAGCAACACAGC 56

RESULT 15
BI165486
LOCUS
DEFINITION
  BI165486 579 bp mRNA linear EST 09-JUL-2001
  RE05402.5prime RE Drosophila melanogaster normalized Embryo pFlc-1
  Drosophila melanogaster cDNA clone RE05402.5 similar to CG3036;
  FBan003036 'transporter' located on: 2L 25B5-25B6; 04/11/2001,
  mRNA sequence.
ACCESSION
  BI165486
VERSION
  BI165486.1 GI:14631292
KEYWORDS
  EST.
SOURCE
  fruit fly.
ORGANISM
  Drosophila melanogaster
  Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
  Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
  Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
  1 (bases 1 to 579)
  
```

**AUTHORS** Stapleton, M., Brokstein, P., Hong, L., Tyler, D., Berman, B., Carlson, J., Champe, M., Chavez, C., Dorsett, V., Farfan, D., Frise, E., George, R., Gonzalez, M., Guarin, H., Harris, N., Li, P., Liao, G., Misra, S., Mungall, C. J., Nunoo, J., Pacleb, J., Paragas, V., Park, S., Phouanavong, S., Wan, K., Yu, C., Lewis, S. E., Celniker, S. and Rubin, G. M.

**TITLE** BDGP/HHMI RE Drosophila EST Project

**JOURNAL** Unpublished (2001)  
**COMMENT** Contact: Stapleton, M.  
 BDGP

Lawrence Berkeley National Lab  
 One Cyclotron Rd, Berkeley, CA 94720, USA  
 Fax: 510 486 6798

Email: [http://www.fruitfly.org/EST\\_est@fruitfly.berkeley.edu](http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu)

hit genomic AE003575: arm:2L [4533649,4860857]

estimated-cyto:25A2-25B8: 04/11/2001

Plate: RE.54 row: A column: 2

High quality sequence stop: 561.

**FEATURES**

source

1. .579

/organism="Drosophila melanogaster"

/db\_xref="taxon:7227"

/clone\_lib="RE Drosophila melanogaster normalized Embryo

pFlc-1"

/sex="male and female"

/dev\_stage="0-24 hours mixed stage embryonic"

/lab\_host="DH5-alpha Tona"

/note="Organ: embryo; Vector: pFlc1; Site:1: XhoI; Site:2:

BamHI; Library was kindly generated by Piero Carninci at

the RIKEN. The library was normalized and excised using

Cre recombinase. Plasmid cDNA library."

BASE COUNT 170 a 131 c 122 g 156 t

**ORIGIN**

Query Match 0.8%; Score 25; DB 13; Length 579;

Best Local Similarity 100.0%; Pred. No. 0.4;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 143 GCAGCAACAGCAGCAGCAACACAGC 167

Db 33 GCAGCAACAGCAGCAGCAGCAACACAGC 57

**RESULT 16**

AQ074074

LOCUS

DEFINITION EP(2)2068 Drosophila melanogaster EP line Drosophila melanogaster

genomic Sequence recovered from Both 5' and 3' ends of P element,

DNA sequence.

ACCESSION AQ074074

VERSION AQ074074.1 GI:3403324

KEYWORDS GSS.

SOURCE fruit fly.

ORGANISM

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 583)

Liao, G.-C., Rehm, E.J. and Rubin, G.M.

Insertion site preferences of the P transposable element in

Drosophila melanogaster

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3347-3351 (2000)

20202638

Contact: Gerald Rubin

Berkeley Drosophila Genome Project

University of California, Berkeley

LSA Building, Berkeley, CA 94720-3200, USA

Fax: 5106439947

Email: [germy@fruitfly.berkeley.edu](mailto:germy@fruitfly.berkeley.edu)

Sequence recovery method was inverse PCR.

Sequence orientation is forward strand relative to 5' end of P

**element**

The P element insertion position is base 155 in the 583 bases. This insertion position refers to the first base of the 8 base target recognition sequence.

Class: transposon-tagged.

**FEATURES**

source

1. .583

/Location/Qualifiers

/organism="Drosophila melanogaster"

/db\_xref="taxon:7227"

/clone\_lib="Drosophila melanogaster EP line"

/note="Inverse PCR was performed on Drosophila

melanogaster strains each of which contains a single EP

transposable element insertion. (The generation of these

insertion strains is described in Rorth P, Szabo K, Bailey

A, Laverty T, Rehm J, Rubin GM, Weigmann K, Milan M, Benes

V, Ansorge W, Cohen SM. 1998. Systematic gain-of-function

genetics in Drosophila. Development 6:1049-1057.) The

resultant fragment for each strain was directly sequenced

to determine the genomic sequence at the site of

insertion. Details of the protocols used can be found at

[http://fruitfly.berkeley.edu/p\\_disrupt/inverse\\_pcr.html](http://fruitfly.berkeley.edu/p_disrupt/inverse_pcr.html)."

BASE COUNT 210 a 148 c 130 g 95 t

**ORIGIN**

Query Match 0.8%; Score 25; DB 17; Length 583;

Best Local Similarity 100.0%; Pred. No. 0.4;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 143 GCAGCAACAGCAGCAGCAACACAGC 167

Db 491 GCAGCAACAGCAGCAGCAGCAACACAGC 515

**RESULT 17**

AQ073492/c

LOCUS

DEFINITION AQ073492

EP(2)2323 Drosophila melanogaster EP line Drosophila melanogaster

genomic Sequence recovered from Both 5' and 3' ends of P element,

DNA sequence.

ACCESSION AQ073492

VERSION AQ073492.1 GI:3403534

KEYWORDS GSS.

SOURCE fruit fly.

ORGANISM

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 586)

Liao, G.-C., Rehm, E.J. and Rubin, G.M.

Insertion site preferences of the P transposable element in

Drosophila melanogaster

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3347-3351 (2000)

20202638

Contact: Gerald Rubin

Berkeley Drosophila Genome Project

University of California, Berkeley

LSA Building, Berkeley, CA 94720-3200, USA

Fax: 5106439947

Email: [germy@fruitfly.berkeley.edu](mailto:germy@fruitfly.berkeley.edu)

Sequence recovery method was inverse PCR.

Sequence orientation is forward strand relative to 5' end of P

element

The P element insertion position is base 116 in the 586 bases. This insertion position refers to the first base of the 8 base target recognition sequence.

Class: transposon-tagged.

**FEATURES**

source

1. .586

/Location/Qualifiers

/organism="Drosophila melanogaster"

/db\_xref="taxon:7227"

/clone\_lib="Drosophila melanogaster EP line"  
 /note="Inverse PCR was performed on Drosophila melanogaster strains each of which contains a single EP transposable element insertion. (The generation of these insertion strains is described in North P, Szabo K, Bailey A, Lavery T, Rehm J, Rubin GM, Weigmann K, Milan M, Benes V, Ansorge W, Cohen SM. 1998. Systematic gain-of-function genetics in Drosophila. Development 6:1049-1057.) The resultant fragment for each strain was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://fruitfly.berkeley.edu/p-disrupt/inverse\_pcr.html."

BASE COUNT 96 a 131 c 151 g 208 t  
 ORIGIN  
 Query Match 0.8%; Score 25; DB 17; Length 586;  
 Best Local Similarity 100.0%; Pred. No. 0.4; Mismatches 0; Indels 0; Gaps 0;  
 Matches 25; Conservative 0;

QY 143 GCAGCAACAGCAGACGACACAGC 167  
 ||||||||||||||||||||||||||||  
 Db 93 GCAGCAACAGCAGACGACACAGC 69

RESULT 18  
 A1387965 590 bp mRNA linear EST 19-APR-2001  
 LOCUS  
 DEFINITION GH18695.5prime GH Drosophila melanogaster head pOT2 Drosophila melanogaster cDNA clone GH18695 5prime, mRNA sequence.

ACCESSION A1387965  
 VERSION  
 KEYWORDS  
 SOURCE EST.  
 ORGANISM fruit fly.

REFERENCE 1 (bases 1 to 590)  
 AUTHORS Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G., Lewis,S. and Rubin,G.M.  
 TITLE BDGP/HMI Drosophila EST Project  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Stapleton, M.

BDGP  
 Lawrence Berkeley National Lab  
 One Cyclotron Rd, Berkeley, CA 94720, USA  
 Fax: 510 486 6798  
 Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu  
 Plate: 186 row: H column: 11  
 High quality sequence stop: 515.  
 Location/Qualifiers  
 1..590  
 /organism="Drosophila melanogaster"  
 /db\_xref="taxon:7227"  
 /clone="GH18695"  
 /clone\_lib="GH Drosophila melanogaster head pOT2"  
 /sex="male and female"  
 /dev\_stage="adult"  
 /lab\_host="DH5 - alpha"  
 /note="Organ: head; Vector: pOT2; Site:1: EcoRI; Site:2: XhoI; Sized fractionated cDNAs were directly ligated into pOT2. Plasmid cDNA library."

BASE COUNT 172 a 135 c 125 g 158 t  
 ORIGIN  
 Query Match 0.8%; Score 25; DB 9; Length 590;  
 Best Local Similarity 100.0%; Pred. No. 0.4; Mismatches 0; Indels 0; Gaps 0;  
 Matches 25; Conservative 0;

QY 143 GCAGCAACAGCAGACGACACAGC 167  
 ||||||||||||||||||||||||||||  
 Db 15 GCAGCAACAGCAGACGACACAGC 39

RESULT 19  
 B1238617 590 bp mRNA linear EST 12-JUL-2001  
 LOCUS  
 DEFINITION RE35121.5prime RE Drosophila melanogaster normalized Embryo pFic-1 Drosophila melanogaster cDNA clone RE35121 5 similar to CG3036: FBan0003036 'transporter' located on: 2L 25B5-25B6;; 05/12/2001, mRNA sequence.

ACCESSION B1238617  
 VERSION B1238617.1 GI:14707119  
 KEYWORDS EST.  
 SOURCE fruit fly.  
 ORGANISM Drosophila melanogaster  
 Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
 REFERENCE 1 (bases 1 to 590)  
 AUTHORS Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson ,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George ,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S., Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S., Phouanavong,S., Wan,K., Yu,C., Lewis,S.E., Ceiniker,S. and Rubin ,G.M.

TITLE BDGP/HMI RE Drosophila EST Project  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Stapleton, M.

BDGP  
 Lawrence Berkeley National Lab  
 One Cyclotron Rd, Berkeley, CA 94720, USA  
 Fax: 510 486 6798  
 Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu  
 hit genomic AE003575: arm:2L [4533649,4860857]  
 estimated-cyto:25A2-25B8: 05/12/2001  
 Plate: RE.351 row: B column: 9  
 High quality sequence stop: 491.  
 Location/Qualifiers  
 1..590  
 /organism="Drosophila melanogaster"  
 /db\_xref="taxon:7227"  
 /clone="RE35121"  
 /clone\_lib="RE Drosophila melanogaster normalized Embryo pFic-1"  
 /sex="male and female"  
 /dev\_stage="0-24 hours mixed stage embryonic"  
 /lab\_host="DH5-alpha TonA"  
 /note="Organ: embryo; Vector: pFic1; Site:1: XhoI; Site:2: BamHI; Library was kindly generated by Piero Carninci at the RIKEN. The library was normalized and excised using Cre recombinase. Plasmid cDNA library."

BASE COUNT 173 a 135 c 123 g 159 t  
 ORIGIN  
 Query Match 0.8%; Score 25; DB 13; Length 590;  
 Best Local Similarity 100.0%; Pred. No. 0.4; Mismatches 25; Conservative 0; Indels 0; Gaps 0;  
 Matches 25; Conservative 0;

QY 143 GCAGCAACAGCAGACGACACAGC 167  
 ||||||||||||||||||||||||||||  
 Db 32 GCAGCAACAGCAGACGACACAGC 56

RESULT 20  
 B1363426 592 bp mRNA linear EST 01-AUG-2001  
 LOCUS  
 DEFINITION RE48231.5prime RE Drosophila melanogaster normalized Embryo pFic-1 Drosophila melanogaster cDNA clone RE48231 5 similar to CG3036: FBan0003036 'transporter' located on: 2L 25B5-25B6;; 05/13/2001, mRNA sequence.

ACCESSION B1363426  
 VERSION B1363426.1 GI:15059454  
 KEYWORDS EST.  
 SOURCE fruit fly.  
 ORGANISM Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

## REFERENCE

1 (bases 1 to 592)  
 AUTHORS Stapleton, M., Brokstein, P., Hong, L., Tyler, D., Berman, B., Carlson, J., Champe, M., Chavez, C., Dorsett, V., Farfan, D., Frise, E., George, R., Gonzalez, M., Guarin, H., Harris, N., Li, P., Liao, G., Misra, S., Mungall, C. J., Nunoo, J., Pacleb, J., Paragas, V., Park, S., Phouanavong, S., Wan, K., Yu, C., Lewis, S. E., Celniker, S. and Rubin, G. M.

TITLE BDGP/HMI RE Drosophila EST Project

JOURNAL Unpublished (2001)

COMMENT Contact: Stapleton, M.

## BDGP

Lawrence Berkeley National Lab  
 One Cyclotron Rd, Berkeley, CA 94720, USA

Fax: 510 486 6798

Email: [http://www.fruitfly.org/EST\\_est@fruitfly.berkeley.edu](http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu)

hit genomic AE003575: arm:2L [4533649,4860857]

estimated-cyto:25A2-25B8: 05/13/2001

Plate: RE.482 row: C column: 7

High quality sequence stop: 472.

## FEATURES

source

1. .592  
 /organism="Drosophila melanogaster"  
 /db\_xref="taxon:7227"  
 /clone="RE48231"  
 /clone\_lib="RE Drosophila melanogaster normalized Embryo pFlc-1"  
 /sex="male and female"  
 /dev\_stage="0-24 hours mixed stage embryonic"  
 /lab\_host="DH5-alpha Tona"  
 /note="Organ: embryo; Vector: pFlc1; Site\_1: XhoI; Site\_2: BamHI. Library was kindly generated by Piero Carninci at the RIKEN. The library was normalized and excised using Cre recombinase. Plasmid cDNA library."  
 BASE COUNT 171 a 137 c 124 g 159 t 1 others  
 ORIGIN

Query Match 0.8%; Score 25; DB 13; Length 592;

Best Local Similarity 100.0%; Pred. No. 0.4;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 143 GCAGCAACAGCAGACACACAGC 167

Db 32 GCAGCAACAGCAGACACACAGC 56

## RESULT 21

BI481998

LOCUS

DEFINITION BI481998 593 bp mRNA linear EST 28-AUG-2001

Drosophila melanogaster cDNA clone RE64636 5 similar to CG3036:

Fban0003036 GO:[transporter (GO:0005215)] located on: 2L 25B5-25B6

:: 05/18/2001, mRNA sequence.

ACCESSION BI481998

VERSION BI481998.1 GI:15321162

KEYWORDS EST.

SOURCE fruit fly.

## ORGANISM

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 593)

## REFERENCE

AUTHORS

Stapleton, M., Brokstein, P., Hong, L., Tyler, D., Berman, B., Carlson, J., Champe, M., Chavez, C., Dorsett, V., Farfan, D., Frise, E., George, R., Gonzalez, M., Guarin, H., Harris, N., Li, P., Liao, G., Misra, S., Mungall, C. J., Nunoo, J., Pacleb, J., Paragas, V., Park, S., Phouanavong, S., Wan, K., Yu, C., Lewis, S. E., Celniker, S. and Rubin, G. M.

TITLE BDGP/HMI RE Drosophila EST Project

JOURNAL Unpublished (2001)

COMMENT Contact: Stapleton, M.

## BDGP

Lawrence Berkeley National Lab  
 One Cyclotron Rd, Berkeley, CA 94720, USA

Fax: 510 486 6798

Email: [http://www.fruitfly.org/EST\\_est@fruitfly.berkeley.edu](http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu)

hit genomic AE003575: arm:2L [4533649,4860857]

estimated-cyto:25A2-25B8: 05/18/2001

Plate: RE.646 row: C column: 12

High quality sequence stop: 561.

## FEATURES

source

1. .593  
 /organism="Drosophila melanogaster"  
 /db\_xref="taxon:7227"  
 /clone="RE64636"  
 /clone\_lib="RE Drosophila melanogaster normalized Embryo pFlc-1"  
 /sex="male and female"  
 /dev\_stage="0-24 hours mixed stage embryonic"  
 /lab\_host="DH5-alpha Tona"  
 /note="Organ: embryo; Vector: pFlc1; Site\_1: XhoI; Site\_2: BamHI. Library was kindly generated by Piero Carninci at the RIKEN. The library was normalized and excised using Cre recombinase. Plasmid cDNA library."  
 BASE COUNT 173 a 136 c 124 g 160 t  
 ORIGIN

Query Match 0.8%; Score 25; DB 13; Length 593;

Best Local Similarity 100.0%; Pred. No. 0.4;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 143 GCAGCAACAGCAGACACACAGC 167

Db 32 GCAGCAACAGCAGACACACAGC 56

## RESULT 22

AI109244

LOCUS

DEFINITION AI109244 602 bp mRNA linear EST 19-APR-2001

melanogaster cDNA clone GH08375 5prime, mRNA sequence.

ACCESSION AI109244.1 GI:3477568

VERSION AI109244.1

KEYWORDS EST.

SOURCE fruit fly.

ORGANISM

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 602)

REFERENCE Harvey, D., Brokstein, P., Hong, L., Evans-Holm, M., Su, C., Tsang, G.,

AUTHORS Lewis, S. and Rubin, G. M.

TITLE BDGP/HMI Drosophila EST Project

JOURNAL Unpublished (2001)

COMMENT Contact: Stapleton, M.

## BDGP

Lawrence Berkeley National Lab

One Cyclotron Rd, Berkeley, CA 94720, USA

Fax: 510 486 6798

Email: [http://www.fruitfly.org/EST\\_est@fruitfly.berkeley.edu](http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu)

Plate: 83 row: G column: 3

High quality sequence stop: 541.

## FEATURES

source

1. .602  
 /organism="Drosophila melanogaster"  
 /db\_xref="taxon:7227"  
 /clone="GH08375"  
 /clone\_lib="GH Drosophila melanogaster head pOT2"  
 /sex="male and female"  
 /dev\_stage="adult"  
 /lab\_host="DH5 - alpha"  
 /note="Organ: head; Vector: pOT2; Site\_1: EcoRI; Site\_2: XhoI; Sized fractionated cDNAs were directly ligated into pOT2. Plasmid cDNA library."

```

BASE COUNT      174 a   140 c   129 g   159 t
ORIGIN
Query Match      0.8%; Score 25; DB 9; Length 602;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 143 GCAGCAACAGCAGCAGCAACACAGC 167
      |||
Db 12 GCAGCAACAGCAGCAGCAACACAGC 36

RESULT 23
LOCUS      AI107819
DEFINITION GH05674.5prime GH Drosophila melanogaster head pOT2 Drosophila
ACCESSION  AI107819
VERSION     AI107819.1 GI:3475472
KEYWORDS   EST.
SOURCE     fruit fly.
ORGANISM   Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE  1 (bases 1 to 616)
AUTHORS   Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
            Lewis,S. and Rubin,G.M.
TITLE     BDGP/HIMI Drosophila EST Project
JOURNAL   Unpublished (2001)
COMMENT   Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            Plate: 56 row: G column: 2
            High quality sequence stop: 547.
            Location/Qualifiers
                Location 1..616
                DB 17 GCAGCAACAGCAGCAGCAACACAGC 41

FEATURES             source
    source
        Query Match      0.8%; Score 25; DB 9; Length 616;
        Best Local Similarity 100.0%; Pred. No. 0.4;
        Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 143 GCAGCAACAGCAGCAGCAACACAGC 167
      |||
Db 17 GCAGCAACAGCAGCAGCAACACAGC 41

RESULT 24
LOCUS      AI063205
DEFINITION GH02745.5prime GH Drosophila melanogaster head pOT2 Drosophila
ACCESSION  AI063205
VERSION     AI063205.1 GI:3339044
KEYWORDS   EST.
SOURCE     fruit fly.
ORGANISM   Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

```

```

REFERENCE
AUTHORS   Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
            Lewis,S. and Rubin,G.M.
TITLE     BDGP/HIMI Drosophila EST Project
JOURNAL   Unpublished (2001)
COMMENT   Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            Plate: 27 row: D column: 9
            High quality sequence stop: 546.
            Location/Qualifiers
                Location 1..625
                ORGANISM="Drosophila melanogaster"
                /db_xref="taxon:7227"
                /clone_lib="GH Drosophila melanogaster head pOT2"
                /sex="male and female"
                /dev_stage="adult"
                /lab_host="DH5 - alpha"
                /note="Organ: head; Vector: pOT2; Site_1: EcoRI; Site_2:
                XhoI; Sized fractionated cDNAs were directly ligated into
                pOT2. Plasmid cDNA library."
BASE COUNT      179 a   151 c   129 g   166 t
ORIGIN
Query Match      0.8%; Score 25; DB 9; Length 625;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 143 GCAGCAACAGCAGCAGCAACACAGC 167
      |||
Db 17 GCAGCAACAGCAGCAGCAACACAGC 41

RESULT 25
LOCUS      AI238498
DEFINITION GH14550.5prime GH Drosophila melanogaster head pOT2 Drosophila
ACCESSION  AI238498
VERSION     AI238498.1 GI:3833356
KEYWORDS   EST.
SOURCE     fruit fly.
ORGANISM   Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE  1 (bases 1 to 625)
AUTHORS   Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
            Lewis,S. and Rubin,G.M.
TITLE     BDGP/HIMI Drosophila EST Project
JOURNAL   Unpublished (2001)
COMMENT   Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            Plate: 145 row: E column: 2
            High quality sequence stop: 518.
            Location/Qualifiers
                Location 1..625
                /organism="Drosophila melanogaster"
                /db_xref="taxon:7227"
                /clone_lib="GH Drosophila melanogaster head pOT2"
                /sex="male and female"
                /dev_stage="adult"

```

```

/lab_host="DH5 - alpha"
/note="Organ: head; Vector: pOT2; Site_1: EcoRI; Site_2:
XhoI; Sized fractionated cDNAs were directly ligated into
pOT2. Plasmid cDNA library."
BASE COUNT      179 a 150 c 132 g 164 t
ORIGIN

Query Match      0.8%; Score 25; DB 9; Length 625;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 143 GCAGCAACAGCAGACGACACACAGC 167
      |||
Db 19 GCAGCAACAGCAGACGACACACAGC 43

RESULT 26
Bil170514
LOCUS
DEFINITION      626 bp mRNA linear EST 09-JUL-2001
Drosophila melanogaster normalized Embryo pFlc-1
Drosophila melanogaster cDNA clone RE12003 5 similar to CG3036;
Fban003036 'transporter' located on: 2L 25B5-25B6;; 04/12/2001,
mRNA sequence.
ACCESSION      Bil170514
VERSION
KEYWORDS
SOURCE
ORGANISM      Drosophila melanogaster
fruit fly.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
AUTHORS      Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S.,
Mungall,C.J., Nunoo,J., Pacieb,J., Paragas,V., Park,S.,
Phouanavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and Rubin
,G.M.
BDGP/HMI RE Drosophila EST Project
Unpublished (2001)
Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
hit genomic AE003575: arm:2L [4533649,4860857]
estimated-cyto:25A2-25B8: 04/12/2001
Plate: RE.120 row: A column: 3
High quality sequence stop: 554.
FEATURES
source
Location/Qualifiers
1..626
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="RE12003"
/pflc-1"
/sex="male and female"
/dev_stage="0-24 hours mixed stage embryonic"
/lab_host="DH5-alpha Tona"
/note="Organ: embryo; Vector: pFlc1; Site_1: XhoI; Site_2:
BamHI; Library was kindly generated by Piero Carninci at
the RIKEN. The library was normalized and excised using
Cre recombinase. Plasmid cDNA library."
BASE COUNT      180 a 146 c 132 g 167 t
ORIGIN

Query Match      0.8%; Score 25; DB 13; Length 626;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 143 GCAGCAACAGCAGACGACACACAGC 167
      |||

```

```

Db 32 GCAGCAACAGCAGACGACACACAGC 56

RESULT 27
A1296017
LOCUS
DEFINITION      628 bp mRNA linear EST 19-APR-2001
Drosophila melanogaster larval-early pupal pOT2
Drosophila melanogaster cDNA clone LP09807 5prime, mRNA sequence.
ACCESSION      A1296017
VERSION
KEYWORDS
SOURCE
ORGANISM      Drosophila melanogaster
fruit fly.
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
AUTHORS      Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
Lewis,S. and Rubin,G.M.
BDGP/HMI Drosophila EST Project
Unpublished (2001)
Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
Plate: 98 row: A column: 7
High quality sequence stop: 500.
FEATURES
source
Location/Qualifiers
1..628
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="LP09807"
/clone_lib="LP Drosophila melanogaster larval-early pupal
pOT2"
/sex="male and female"
/dev_stage="larvae-pupae"
/lab_host="DH5-alpha"
/note="Organ: whole body; Vector: pOT2; Site_1: EcoRI;
Site_2: XhoI; Sized fractionated cDNAs were directly
ligated into pOT2. Plasmid cDNA library."
BASE COUNT      179 a 149 c 134 g 166 t
ORIGIN

Query Match      0.8%; Score 25; DB 9; Length 628;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 143 GCAGCAACAGCAGACGACACACAGC 167
      |||
Db 3 GCAGCAACAGCAGACGACACACAGC 27

RESULT 28
Bil169402
LOCUS
DEFINITION      633 bp mRNA linear EST 09-JUL-2001
Drosophila melanogaster normalized Embryo pFlc-1
Fban003036 'transporter' located on: 2L 25B5-25B6;; 04/11/2001,
mRNA sequence.
ACCESSION      Bil169402
VERSION
KEYWORDS
SOURCE
ORGANISM      Drosophila melanogaster
fruit fly.
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
AUTHORS      Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George

```

R., Gonzalez, M., Guarin, H., Harris, N., Li, P., Liao, G., Misra, S., Mungall, C. J., Nunoo, J., Paclebe, J., Paragas, V., Park, S., Phouanavong, S., Wan, K., Yu, C., Lewis, S. E., Celniker, S. and Rubin, G. M.

TITLE BDGP/HHMI RE Drosophila EST Project  
JOURNAL Unpublished (2001)  
COMMENT Contact: Stapleton, M.  
BDGP

Lawrence Berkeley National Lab  
One Cyclotron Rd, Berkeley, CA 94720, USA  
Fax: 510 486 6798

Email: <http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu>  
hit genomic AE003575: arm:2L [4533649,4860857]  
estimated-cyto:25A2-25B8: 04/11/2001  
Plate: RE.105 row: F column: 6  
High quality sequence stop: 560.

#### FEATURES

source  
1. .633  
Location/Qualifiers

/organism="Drosophila melanogaster"  
/db\_xref="taxon:7227"  
/clone="RE10566"  
/clone\_lib="RE Drosophila melanogaster normalized Embryo  
pFlc-1"

/sex="male and female"

/dev\_stage="0-24 hours mixed stage embryonic"

/lab\_host="DH5-alpha Tona"

/note="Organ: embryo; Vector: pFlc1; Site:1; XhoI: Site:2;  
BamHI: Library was kindly generated by Piero Carninci at  
the RIKEN. The library was normalized and excised using  
Cre recombinase. Plasmid cDNA library."

182 a 149 c 132 g 169 t 1 others

BASE COUNT  
ORIGIN

Query Match 0.8%; Score 25; DB 13; Length 633;

Best Local Similarity 100.0%; Pred. No. 0.4;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 143 GCAGCAACAGCAGACAAACACAGC 167

Db 32 GCAGCAACAGCAGACAAACACAGC 56

RESULT 29

AI386994

LOCUS

DEFINITION GH17396.5prime GH Drosophila melanogaster head pOT2 Drosophila melanogaster cDNA clone GH17396 5prime, mRNA sequence.

ACCESSION

AI386994

VERSION

AI386994.1

KEYWORDS

EST.

SOURCE

fruit fly.

ORGANISM

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 639)

Harvey, D., Brokstein, P., Hong, L., Evans-Holm, M., Su, C., Tsang, G.,

Lewis, S. and Rubin, G.M.

BDGP/HHMI Drosophila EST Project

Unpublished (2001)

Contact: Stapleton, M.

BDGP

Lawrence Berkeley National Lab

One Cyclotron Rd, Berkeley, CA 94720, USA

Fax: 510 486 6798

Email: <http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu>

Plate: 173 row: H column: 12

High quality sequence stop: 546.

FEATURES

source

1. .633

Location/Qualifiers

/organism="Drosophila melanogaster"

/db\_xref="taxon:7227"

/clone="GH17396"

/clone\_lib="GH Drosophila melanogaster head pOT2"  
/sex="male and female"  
/dev\_stage="adult"  
/lab\_host="DH5 - alpha"  
/note="Organ: head; Vector: pOT2; Site:1; EcoRI: Site:2;  
XhoI: Sized fractionated cDNAs were directly ligated into  
pOT2. Plasmid cDNA library."

BASE COUNT 181 a 153 c 136 g 169 t

ORIGIN

Query Match 0.8%; Score 25; DB 9; Length 639;

Best Local Similarity 100.0%; Pred. No. 0.4;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 143 GCAGCAACAGCAGACAAACACAGC 167

Db 17 GCAGCAACAGCAGACAAACACAGC 41

RESULT 30

AI388097

LOCUS

DEFINITION AI388097 655 bp mRNA linear EST 19-APR-2001  
melanogaster cDNA clone GH18879 5prime, mRNA sequence.

ACCESSION

AI388097

VERSION

AI388097.1

KEYWORDS

EST.

SOURCE

fruit fly.

ORGANISM

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 655)

Harvey, D., Brokstein, P., Hong, L., Evans-Holm, M., Su, C., Tsang, G.,

Lewis, S. and Rubin, G.M.

BDGP/HHMI Drosophila EST Project

Unpublished (2001)

Contact: Stapleton, M.

BDGP

Lawrence Berkeley National Lab

One Cyclotron Rd, Berkeley, CA 94720, USA

Fax: 510 486 6798

Email: <http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu>

Plate: 188 row: G column: 7

High quality sequence stop: 536.

FEATURES

source

1. .655

Location/Qualifiers

/organism="Drosophila melanogaster"

/db\_xref="taxon:7227"

/clone="GH18879"

/clone\_lib="GH Drosophila melanogaster head pOT2"

/sex="male and female"

/dev\_stage="adult"

/lab\_host="DH5 - alpha"

/note="Organ: head; Vector: pOT2; Site:1; EcoRI: Site:2;

XhoI: Sized fractionated cDNAs were directly ligated into

pOT2. Plasmid cDNA library."

187 a 159 c 138 g 171 t

BASE COUNT

ORIGIN

Query Match

Best Local Similarity 100.0%; Pred. No. 0.4;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 143 GCAGCAACAGCAGACAAACACAGC 167

Db 15 GCAGCAACAGCAGACAAACACAGC 39

RESULT 31

AI107422

LOCUS

DEFINITION GH05102.5prime GH Drosophila melanogaster head pOT2 Drosophila melanogaster cDNA clone GH05102 5prime, mRNA linear EST 19-APR-2001

```

melanogaster cDNA clone GH05102 5prime, mRNA sequence.
ACCESSION      AI107422
VERSION        AI107422.1  GI:3475075
KEYWORDS       EST.
SOURCE         fruit fly.
ORGANISM       Drosophila melanogaster
               Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
               Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
               Ephydroidea; Drosophilidae; Drosophila.
REFERENCE      1 (bases 1 to 664)
AUTHORS       Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
               Lewis,S. and Rubin,G.M.
TITLE         BDGP/HHMI Drosophila EST Project
JOURNAL        Unpublished (2001)
COMMENT       Other ESTs: GH05102.3prime
               Contact: Stapleton, M.
               BDGP
               Lawrence Berkeley National Lab
               One Cyclotron Rd, Berkeley, CA 94720, USA
               Fax: 510 486 6798
               Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
               Plate: 51 row: A column: 2
               High quality sequence stop: 555.

FEATURES             Location/Qualifiers
     source            1..664
                       /organism="Drosophila melanogaster"
                       /db_xref="taxon:7227"
                       /clone_lib="GH Drosophila melanogaster head pOT2"
                       /sex="male and female"
                       /dev_stage="adult"
                       /lab_host="DH5 - alpha"
                       /note="Organ: head; Vector: pOT2; Site_1: EcoRI; Site_2:
                       XhoI; Sized fractionated cDNAs were directly ligated into
                       pOT2. Plasmid cDNA library."
BASE COUNT          189 a 161 c 138 g 175 t
ORIGIN
1..664
+
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="GH05102"
/clone_lib="GH Drosophila melanogaster head pOT2"
/sex="male and female"
/dev_stage="adult"
/lab_host="DH5 - alpha"
/note="Organ: head; Vector: pOT2; Site_1: EcoRI; Site_2:
XhoI; Sized fractionated cDNAs were directly ligated into
pOT2. Plasmid cDNA library."
1 others
BASE COUNT          189 a 161 c 138 g 175 t
ORIGIN
1..664
+
Query Match          0.8%; Score 25; DB 9; Length 664;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 143 GCAGCAACAGCAGCAGCAACACAGC 167
+++++
Db 25 GCAGCAACAGCAGCAGCAACACAGC 49
+++++

RESULT 32
AI292862
LOCUS            AI292862
DEFINITION      GH15875.5prime GH Drosophila melanogaster head pOT2 Drosophila
melanogaster cDNA clone GH15875 5prime, mRNA sequence.
ACCESSION      AI292862
VERSION        AI292862.1  GI:3942269
KEYWORDS       EST.
SOURCE         fruit fly.
ORGANISM       Drosophila melanogaster
               Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
               Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
               Ephydroidea; Drosophilidae; Drosophila.
REFERENCE      1 (bases 1 to 664)
AUTHORS       Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
               Lewis,S. and Rubin,G.M.
TITLE         BDGP/HHMI Drosophila EST Project
JOURNAL        Unpublished (2001)
COMMENT       Contact: Stapleton, M.
               BDGP
               Lawrence Berkeley National Lab
               One Cyclotron Rd, Berkeley, CA 94720, USA
               Fax: 510 486 6798
               Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
               Plate: 158 row: G column: 3
               High quality sequence stop: 509.

melanogaster cDNA clone GH05102 5prime, mRNA sequence.
ACCESSION      AI107422
VERSION        AI107422.1  GI:3475075
KEYWORDS       EST.
SOURCE         fruit fly.
ORGANISM       Drosophila melanogaster
               Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
               Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
               Ephydroidea; Drosophilidae; Drosophila.
REFERENCE      1 (bases 1 to 664)
AUTHORS       Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
               Lewis,S. and Rubin,G.M.
TITLE         BDGP/HHMI Drosophila EST Project
JOURNAL        Unpublished (2001)
COMMENT       Contact: Stapleton, M.
               BDGP
               Lawrence Berkeley National Lab
               One Cyclotron Rd, Berkeley, CA 94720, USA
               Fax: 510 486 6798
               Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
               Plate: 158 row: G column: 3
               High quality sequence stop: 509.

```

```

FEATURES             Location/Qualifiers
     source            1..664
                       /organism="Drosophila melanogaster"
                       /db_xref="taxon:7227"
                       /clone_lib="GH Drosophila melanogaster head pOT2"
                       /sex="male and female"
                       /dev_stage="adult"
                       /lab_host="DH5 - alpha"
                       /note="Organ: head; Vector: pOT2; Site_1: EcoRI; Site_2:
                       XhoI; Sized fractionated cDNAs were directly ligated into
                       pOT2. Plasmid cDNA library."
BASE COUNT          189 a 160 c 141 g 174 t
ORIGIN
1..664
+
Query Match          0.8%; Score 25; DB 9; Length 664;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 143 GCAGCAACAGCAGCAGCAACACAGC 167
+++++
Db 17 GCAGCAACAGCAGCAGCAACACAGC 41
+++++

RESULT 33
AI109291
LOCUS            AI109291
DEFINITION      AT109291
                 666 bp mRNA linear EST 19-APR-2001
                 GH08434.5prime GH Drosophila melanogaster head pOT2 Drosophila
                 melanogaster cDNA clone GH08434 5prime, mRNA sequence.
ACCESSION      AI109291
VERSION        AI109291.1  GI:3477615
KEYWORDS       EST.
SOURCE         fruit fly.
ORGANISM       Drosophila melanogaster
               Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
               Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
               Ephydroidea; Drosophilidae; Drosophila.
REFERENCE      1 (bases 1 to 666)
AUTHORS       Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
               Lewis,S. and Rubin,G.M.
TITLE         BDGP/HHMI Drosophila EST Project
JOURNAL        Unpublished (2001)
COMMENT       Contact: Stapleton, M.
               BDGP
               Lawrence Berkeley National Lab
               One Cyclotron Rd, Berkeley, CA 94720, USA
               Fax: 510 486 6798
               Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
               Plate: 84 row: C column: 10
               High quality sequence stop: 555.

FEATURES             Location/Qualifiers
     source            1..666
                       /organism="Drosophila melanogaster"
                       /db_xref="taxon:7227"
                       /clone_lib="GH Drosophila melanogaster head pOT2"
                       /sex="male and female"
                       /dev_stage="adult"
                       /lab_host="DH5 - alpha"
                       /note="Organ: head; Vector: pOT2; Site_1: EcoRI; Site_2:
                       XhoI; Sized fractionated cDNAs were directly ligated into
                       pOT2. Plasmid cDNA library."
BASE COUNT          189 a 162 c 140 g 175 t
ORIGIN
1..666
+
Query Match          0.8%; Score 25; DB 9; Length 666;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 143 GCAGCAACAGCAGCAGCAACACAGC 167
+++++
Db 25 GCAGCAACAGCAGCAGCAACACAGC 49
+++++

```



## RESULT 34

BI368317

## LOCUS

DEFINITION RE54268.5prime RE Drosophila melanogaster normalized Embryo pF1c-1 Drosophila melanogaster cDNA clone RE54268 5 similar to CG3036: FBan0003036 'transporter' located on: 2L 25B5-25B6:: 05/14/2001, mRNA sequence.

## ACCESSION

BI368317

## VERSION

BI368317.1

## KEYWORDS

EST.

## SOURCE

fruit fly.

## ORGANISM

Drosophila melanogaster

## REFERENCE

AUTHORS

1 (bases 1 to 677) Stapleton, M., Brokstein, P., Hong, L., Tyler, D., Berman, B., Carlson, J., Champe, M., Chavez, C., Dorsett, V., Farfan, D., Frise, E., George, R., Gonzalez, M., Guarin, H., Harris, N., Li, P., Liao, G., Misra, S., Mungall, C.J., Nunoo, J., Pacleb, J., Paragas, V., Park, S., Phouanavong, S., Wan, K., Yu, C., Lewis, S.E., Celniker, S. and Rubin, G.M.

## TITLE

BDGP/HIMI RE Drosophila EST Project

## JOURNAL

Unpublished (2001)

## COMMENT

Contact: Stapleton, M.

## BDGP

Lawrence Berkeley National Lab

One Cyclotron Rd, Berkeley, CA 94720, USA

Fax: 510 486 6798

Email: <http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu>

hit genomic AE003575: arm:2L [4533649,4860857]

estimated-cyto:25A2-25B8: 05/14/2001

Plate: RE 542 row: F column: 8

High quality sequence stop: 559.

## FEATURES

source

1..677

Location/Qualifiers

/organism="Drosophila melanogaster"

/db\_xref="taxon:7227"

/clone="RE54268"

/clone\_lib="RE Drosophila melanogaster normalized Embryo

pF1c-1"

/sex="male and female"

/dev\_stage="0-24 hours mixed stage embryonic"

/lab\_host="DH5-alpha Tona"

/note="Organ: embryo; Vector: pF1c1; Site\_1: XhoI; Site\_2:

BamHI; Library was kindly generated by Piero Carninci at

the RIKEN. The library was normalized and excised using

Cre recombinase. Plasmid cDNA library."

BASE COUNT 190 a 163 c 145 g 179 t

## ORIGIN

1..677

Query Match 0.8%; Score 25; DB 13; Length 677;

Best Local Similarity 100.0%; Pred. No. 0.4;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 143 GCAGCAACAGCAGCAACACAGC 167

|||||

Db 30 GCAGCAACAGCAGCAACACAGC 54

## RESULT 35

BI374631

## LOCUS

DEFINITION RE62269.5prime RE Drosophila melanogaster normalized Embryo pF1c-1 Drosophila melanogaster cDNA clone RE62269 5 similar to CG3036: FBan0003036 GO:[transporter (GO:0005215)] located on: 2L 25B5-25B6:: 05/16/2001, mRNA sequence.

## ACCESSION

BI374631

## VERSION

BI374631.1

## KEYWORDS

EST.

## SOURCE

fruit fly.

## ORGANISM

Drosophila melanogaster

## REFERENCE

AUTHORS

1 (bases 1 to 686) Stapleton, M., Brokstein, P., Hong, L., Tyler, D., Berman, B., Carlson, J., Champe, M., Chavez, C., Dorsett, V., Farfan, D., Frise, E., George, R., Gonzalez, M., Guarin, H., Harris, N., Li, P., Liao, G., Misra, S., Mungall, C.J., Nunoo, J., Pacleb, J., Paragas, V., Park, S., Phouanavong, S., Wan, K., Yu, C., Lewis, S.E., Celniker, S. and Rubin, G.M.

## TITLE

BDGP/HIMI RE Drosophila EST Project

## JOURNAL

Unpublished (2001)

## COMMENT

Contact: Stapleton, M.

## BDGP

Lawrence Berkeley National Lab

One Cyclotron Rd, Berkeley, CA 94720, USA

Fax: 510 486 6798

Email: <http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu>

hit genomic AE003575: arm:2L [4533649,4860857]

estimated-cyto:25A2-25B8: 05/16/2001

Plate: RE 622 row: F column: 9

High quality sequence stop: 637.

## FEATURES

source

1..686

Location/Qualifiers

/organism="Drosophila melanogaster"

/db\_xref="taxon:7227"

/clone="RE62269"

/clone\_lib="RE Drosophila melanogaster normalized Embryo

pF1c-1"

/sex="male and female"

/dev\_stage="0-24 hours mixed stage embryonic"

/lab\_host="DH5-alpha Tona"

/note="Organ: embryo; Vector: pF1c1; Site\_1: XhoI; Site\_2:

BamHI; Library was kindly generated by Piero Carninci at

the RIKEN. The library was normalized and excised using

Cre recombinase. Plasmid cDNA library."

BASE COUNT 194 a 165 c 146 g 181 t

## ORIGIN

1..686

Query Match 0.8%; Score 25; DB 13; Length 686;

Best Local Similarity 100.0%; Pred. No. 0.4;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 143 GCAGCAACAGCAGCAACACAGC 167

|||||

Db 30 GCAGCAACAGCAGCAACACAGC 54

## RESULT 36

BI368263

## LOCUS

DEFINITION

RE54204.5prime RE Drosophila melanogaster normalized Embryo pF1c-1 Drosophila melanogaster cDNA clone RE54204 5 similar to CG3036: FBan0003036 'transporter' located on: 2L 25B5-25B6:: 05/14/2001, mRNA sequence.

## ACCESSION

BI368263

## VERSION

BI368263.1

## KEYWORDS

EST.

## SOURCE

fruit fly.

## ORGANISM

Drosophila melanogaster

## REFERENCE

AUTHORS

1 (bases 1 to 687) Stapleton, M., Brokstein, P., Hong, L., Tyler, D., Berman, B., Carlson, J., Champe, M., Chavez, C., Dorsett, V., Farfan, D., Frise, E., George, R., Gonzalez, M., Guarin, H., Harris, N., Li, P., Liao, G., Misra, S., Mungall, C.J., Nunoo, J., Pacleb, J., Paragas, V., Park, S., Phouanavong, S., Wan, K., Yu, C., Lewis, S.E., Celniker, S. and Rubin, G.M.

## TITLE

BDGP/HIMI RE Drosophila EST Project

## JOURNAL

Unpublished (2001)

## COMMENT

Contact: Stapleton, M.

```

BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
hit genomic AE003575: arm:2L [4533649,4860857]
estimated-cyto:25A2-25B8: 05/14/2001
Plate: RE.542 row: A column: 4
High quality sequence stop: 560.
Location/Qualifiers
1..687
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="RE54204"
/clone_lib="RE Drosophila melanogaster normalized Embryo
pFlc-1"
/sex="male and female"
/dev_stage="0-24 hours mixed stage embryonic"
/lab_host="DH5-alpha Tona"
/note="Organ: embryo; Vector: pFlc1; Site_1: XhoI; Site_2:
BamHI; Library was kindly generated by Piero Carninci at
the RIKEN. The library was normalized and excised using
Cre recombinase. Plasmid cDNA library."
BASE COUNT      194 a 167 c 145 g 181 t
ORIGIN
Query Match      0.8%; Score 25; DB 13; Length 687;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 143 GCAGCAACAGCAGCAGCAACACAGC 167
|||||
Db 32 GCAGCAACAGCAGCAGCAACACAGC 56

RESULT 37
A1517363
LOCUS      734 bp mRNA linear EST 19-APR-2001
DEFINITION GH28085.5prime GH Drosophila melanogaster head pOT2 Drosophila
melanogaster cDNA clone GH28085 5prime, mRNA sequence.
ACCESSION  A1517363
VERSION    A1517363.1 GI:4420463
KEYWORDS   EST.
SOURCE     fruit fly.
ORGANISM   Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE  Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
AUTHORS   Lewis,S. and Rubin,G.M.
TITLE     BDGP/HMI Drosophila EST Project
JOURNAL   Unpublished (2001)
COMMENT   Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
Plate: 280 row: H column: 1
High quality sequence stop: 541.
Location/Qualifiers
1..734
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="GH28085"
/clone_lib="GH Drosophila melanogaster head pOT2"
/sex="male and female"
/dev_stage="adult"
/lab_host="DH5 - alpha"
/note="Organ: head; Vector: pOT2; Site_1: EcoRI; Site_2:
XhoI; Sized fractionated cDNAs were directly ligated into
pOT2. Plasmid cDNA library."

```

```

BASE COUNT      206 a 176 c 168 g 184 t
ORIGIN
Query Match      0.8%; Score 25; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 143 GCAGCAACAGCAGCAGCAACACAGC 167
|||||
Db 12 GCAGCAACAGCAGCAGCAACACAGC 36

RESULT 38
AZ628307/c
LOCUS      496 bp DNA linear GSS 13-DEC-2000
DEFINITION 1M0480018F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0480018 F, DNA sequence.
ACCESSION  AZ628307
VERSION    AZ628307.1 GI:11750497
KEYWORDS   GSS.
SOURCE     house mouse.
ORGANISM   Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
AUTHORS   Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
JOURNAL   Unpublished (2000)
COMMENT   Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0480 row: O column: 18
Seq primer: CGTTGTAACAGCAGCGCCAGT
Class: plasmid ends
High quality sequence stop: 496.
Location/Qualifiers
1..496
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0480018"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pWD42 (gi14732114|gbAF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptored mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

```

BASE COUNT 79 a 194 c 70 g 153 t

Query Match 0.7%; Score 21; DB 17; Length 496;  
Best Local Similarity 100.0%; Pred. No. 43;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2454 CGGTGTTGAAGGCAAAATGCG 2474  
|||||

Db 28 CGGTGTTGAAGGCAAAATGCG 8

# RESULT 39

BI468065/c

## LOCUS

DEFINITION EST00475 Atlantic salmon Lambda Zap Express liver cDNA library  
protein variant 1, mRNA sequence.

## ACCESSION

VERSION BI468065

## KEYWORDS

SOURCE Atlantic salmon.

## ORGANISM

Salmo salar  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.

## REFERENCE

AUTHORS Martin, S.A., Caplice, N.C., Davey, G.C. and Powell, R.

TITLE An expressed sequence tag based list of genes expressed in the

JOURNAL liver of Atlantic salmon (Salmo salar)

## COMMENT

Unpublished (2001)

Contact: Martin SA

Department of Microbiology

National University of Ireland Galway (NUIG)

National University of Ireland Galway, Galway, Ireland

Tel: 00353 91 524411 2254

Fax: 00353 91 525700

Email: sarah.martine@nui-galway.ie

Insert Length: 593 Std Error: 0.00

Plate: Liver rare plate 5 row: g column: 04

Seq primer: M13 reverse primer - caggaacagctatgacc

High quality sequence stop: 593

## FEATURES

source

1..593

/organism="Salmo salar"

/db\_xref="taxon:8030"

/clone\_lib="LRR5-g04"

/clone\_lib="Atlantic salmon Lambda Zap Express liver cDNA

library"

/tissue\_type="liver"

/dev\_stage="adult"

/note="organ: liver; Vector: Lambda ZAP Express; Site\_1:  
EcoRI; Site\_2: XhoI; An Atlantic salmon liver cDNA library  
was constructed using the Lambda Zap Express/Gigapack  
cloning kit (Stratagene cloning systems). cDNA synthesis  
was carried out using an oligo-(dT) primer for the  
reverse transcription of 5ug of mRNA and the library was  
constructed by directional cloning EcoRI-XhoI based on  
manufacturers instructions. An insert:vector ligation  
ratio of 1:5 was chosen as most optimum. The lambda  
library was packaged with Gigapack III gold packaging  
extracts and plated on the E. coli cell line XL1-Blue  
MRF'."

BASE COUNT 152 a 182 c 141 g 118 t

## ORIGIN

Query Match

Best Local Similarity 0.7%; Score 21; DB 13; Length 593;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2710 CATGCGAAGGCGCGTCAAC 2730

|||||

Db 352 CATGCGAAGGCGCGTCAAC 332

# RESULT 40

AI238153

## LOCUS

DEFINITION AI238153 727 bp mRNA linear EST 19-APR-2001  
melanogaster cDNA clone GH14154 5prime, mRNA sequence.

## ACCESSION

VERSION AI238153

## KEYWORDS

SOURCE EST.

## ORGANISM

Drosophila melanogaster  
fruit fly.  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.

## REFERENCE

AUTHORS Harvey, D., Brokstein, P., Hong, L., Evans-Holm, M., Su, C., Tsang, G.,  
Lewis, S. and Rubin, G.M.

TITLE BDGP/HHMI Drosophila EST Project

JOURNAL Unpublished (2001)

COMMENT Contact: Stapleton, M.

## ORGANISM

BDGP

Lawrence Berkeley National Lab

One Cyclotron Rd, Berkeley, CA 94720, USA

Fax: 510 486 6798

Email: <http://www.fruitfly.org/EST>, [est@fruitfly.berkeley.edu](mailto:est@fruitfly.berkeley.edu)

Plate: 141 row: E column: 6

High quality sequence stop: 597.

## FEATURES

source

1..727

/organism="Drosophila melanogaster"

/db\_xref="taxon:7227"

/clone="GH14154"

/clone\_lib="GH Drosophila melanogaster head pOT2"

/sex="male and female"

/dev\_stage="adult"

/lab\_host="DH5 - alpha"

/note="Organ: head; Vector: pOT2; Site\_1: EcoRI; Site\_2:  
XhoI; Sized fractionated cDNAs were directly ligated into  
pOT2. Plasmid cDNA library."

BASE COUNT 201 a 173 c 164 g 189 t

## ORIGIN

Query Match

Best Local Similarity 0.7%; Score 21; DB 9; Length 727;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 147 CAACAGCAGACGACACACAGC 167

|||||

Db 1 CAACAGCAGACGACACACAGC 21

Search completed: January 27, 2003, 23:05:37

Job time : 2950 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 27, 2003, 20:57:34 ; Search time 81 Seconds  
(without alignments)  
12130.770 Million cell updates/sec

Title: US-09-830-433A-7

Perfect score: 3204

Sequence: 1 atggaacagaccccaacatt.....gcgtaggtaacggtctga 3204

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 441362 seqs, 153338381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Issued\_Patents\_NA.\*

- 1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*
- 2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*
- 3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*
- 4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*
- 5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq.\*
- 6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	20	0.6	735	3	US-08-750-145A-23
C 2	20	0.6	735	3	US-08-975-698A-27
C 3	20	0.6	735	4	US-09-417-090-27
C 4	20	0.6	735	4	US-09-727-578-27
C 5	18	0.6	60	1	US-08-219-012-74
C 6	18	0.6	60	4	US-08-687-421-262
C 7	18	0.6	692	4	US-09-221-017B-718
C 8	18	0.6	1689	4	US-09-276-531-101
C 9	17	0.5	77	1	US-08-248-474-87
C 10	17	0.5	77	3	US-08-756-849-87
C 11	17	0.5	117	4	US-08-906-156A-27
C 12	17	0.5	322	4	US-08-906-156A-28
C 13	17	0.5	362	4	US-09-641-638-222
C 14	17	0.5	362	4	US-09-641-638-223
C 15	17	0.5	362	4	US-09-641-638-224
C 16	17	0.5	362	4	US-09-641-638-225
C 17	17	0.5	429	4	US-09-060-756-653
C 18	17	0.5	500	4	US-08-642-274D-18
C 19	17	0.5	500	4	US-08-952-014C-18
C 20	17	0.5	671	4	US-08-998-416-826
C 21	17	0.5	800	5	PCT-US95-04801-4
C 22	17	0.5	871	4	US-08-906-156A-13
C 23	17	0.5	871	4	US-08-906-156A-58
C 24	17	0.5	925	3	US-09-267-031-5
C 25	17	0.5	981	2	US-08-770-565-1
C 26	17	0.5	981	2	US-08-710-249-5
C 27	17	0.5	981	2	US-08-833-377-1

Sequence 2, Appli	981	2	US-08-714-482-2	17	C 28
Sequence 22, Appli	981	3	US-08-838-545-22	17	C 29
Sequence 22, Appli	981	4	US-09-349-532-22	17	C 30
Sequence 5, Appli	981	4	US-09-220-157A-5	17	C 31
Sequence 30, Appli	1035	1	US-07-601-094-30	17	C 32
Sequence 30, Appli	1035	1	US-08-012-735-30	17	C 33
Sequence 8, Appli	1257	4	US-08-791-115B-8	17	C 34
Sequence 5, Appli	1440	2	US-08-224-483-5	17	C 35
Sequence 11, Appli	1773	4	US-08-906-156A-11	17	C 36
Sequence 7, Appli	1773	4	US-09-310-363C-7	17	C 37
Sequence 1, Appli	1801	5	PCT-US95-02455-1	17	C 38
Sequence 3, Appli	1910	4	US-09-593-711A-3	17	C 39
Sequence 1, Appli	1914	1	US-07-601-094-1	17	C 40
Sequence 1, Appli	1914	1	US-08-012-735-1	17	C 41
Sequence 3, Appli	1962	4	US-08-791-115B-3	17	C 42
Patent No. 5349058	1971	6	5349058-1	17	C 43
Sequence 1, Appli	2139	3	US-09-026-343-1	17	C 44
Sequence 1, Appli	2139	4	US-09-362-871-1	17	C 45
Sequence 7, Appli	2406	1	US-08-396-479B-7	17	C 46
Sequence 7, Appli	2406	1	US-08-818-823-7	17	C 47
Sequence 3, Appli	2420	1	US-08-330-123A-3	17	C 48
Sequence 3, Appli	2420	4	US-09-580-517-3	17	C 49
Sequence 1, Appli	2425	2	US-08-485-778-1	17	C 50
Sequence 1, Appli	2425	3	US-08-520-550A-1	17	C 51
Sequence 3, Appli	2426	1	US-08-482-115B-3	17	C 52
Sequence 3, Appli	2426	2	US-08-660-678A-3	17	C 53
Sequence 3, Appli	2426	2	US-08-472-802C-4	17	C 54
Sequence 1, Appli	2426	2	US-08-714-482-1	17	C 55
Sequence 3, Appli	2426	3	US-08-998-443-3	17	C 56
Sequence 3, Appli	2426	4	US-09-060-523-3	17	C 57
Sequence 9, Appli	2647	1	US-08-396-479B-9	17	C 58
Sequence 9, Appli	2647	1	US-08-818-823-9	17	C 59
Sequence 1, Appli	2759	4	US-09-144-367-1	17	C 60
Sequence 3, Appli	2759	2	US-08-937-931-3	17	C 61
Sequence 3, Appli	2796	4	US-09-285-502-3	17	C 62
Sequence 3, Appli	2796	4	US-09-709-126-3	17	C 63
Sequence 3, Appli	2796	4	US-09-871-385A-3	17	C 64
Sequence 31, Appli	2811	1	US-08-040-548-31	17	C 65
Sequence 31, Appli	2811	1	US-08-466-344-31	17	C 66
Patent No. 5206152	2817	6	5206152-6	17	C 67
Sequence 2, Appli	3160	4	US-08-791-115B-2	17	C 68
Sequence 1, Appli	3330	1	US-08-149-103-1	17	C 69
Sequence 1, Appli	3330	1	US-08-451-883-1	17	C 70
Sequence 1, Appli	3349	2	US-08-920-234-1	17	C 71
Sequence 3, Appli	3410	4	US-09-527-154-3	17	C 72
Sequence 1, Appli	3656	1	US-08-393-734-1	17	C 73
Sequence 1, Appli	3656	4	US-08-894-489-1	17	C 74
Sequence 18, Appli	3695	4	US-09-453-702B-18	17	C 75
Sequence 11, Appli	3969	1	US-08-396-479B-11	17	C 76
Sequence 11, Appli	3969	1	US-08-818-823-11	17	C 77
Sequence 40, Appli	4010	5	PCT-US94-07297-40	17	C 78
Sequence 1, Appli	4534	2	US-08-935-450-1	17	C 79
Sequence 25, Appli	4534	4	US-09-166-350-25	17	C 80
Sequence 9, Appli	8147	4	US-09-514-247A-9	17	C 81
Sequence 3, Appli	9592	1	US-08-393-734-3	17	C 82
Sequence 3, Appli	9592	4	US-08-894-489-3	17	C 83
Sequence 1, Appli	10747	2	US-08-147-777-1	17	C 84
Sequence 1, Appli	10747	3	US-08-452-872-1	17	C 85
Sequence 1, Appli	10747	5	PCT-US93-03985-1	17	C 86
Sequence 3, Appli	24979	2	US-08-147-777-3	17	C 87
Sequence 3, Appli	24979	3	US-08-452-872-3	17	C 88
Sequence 3, Appli	24979	5	PCT-US93-03985-3	17	C 89
Sequence 2, Appli	4403765	4	US-09-103-840A-2	17	C 90
Sequence 2, Appli	4403765	4	US-09-103-840A-2	17	C 91
Sequence 1, Appli	4411529	4	US-09-103-840A-1	17	C 92
Sequence 1, Appli	4411529	4	US-09-103-840A-1	17	C 93
Sequence 10, Appli	18	2	US-09-205-860-10	16	C 94
Sequence 13, Appli	20	4	US-09-517-584A-13	16	C 95
Sequence 41, Appli	322	4	US-09-117-121-41	16	C 96
Sequence 7, Appli	322	4	US-09-344-523-7	16	C 97
Sequence 96, Appli	386	3	US-08-804-439A-96	16	C 98
Sequence 96, Appli	386	3	US-08-720-229-96	16	C 99
Sequence 3, Appli	390	4	US-08-930-264-3	16	C 100

101	16	0.5	390	4	US-08-930-264-5	Sequence 5, Appli	174	16	0.5	1247	1	US-08-278-729A-32	Sequence 32, Appl
102	16	0.5	390	4	US-08-930-264-7	Sequence 7, Appli	175	16	0.5	1247	1	US-08-155-343A-32	Sequence 32, Appl
103	16	0.5	390	4	US-08-930-264-9	Sequence 9, Appli	176	16	0.5	1247	1	US-08-406-672-32	Sequence 32, Appl
104	16	0.5	390	4	US-08-930-264-11	Sequence 11, Appl	177	16	0.5	1247	1	US-08-643-563A-32	Sequence 32, Appl
105	16	0.5	390	4	US-08-930-264-13	Sequence 13, Appl	178	16	0.5	1247	1	US-08-643-763A-32	Sequence 32, Appl
106	16	0.5	390	4	US-08-930-264-15	Sequence 15, Appl	179	16	0.5	1247	1	US-08-462-623-32	Sequence 32, Appl
107	16	0.5	390	4	US-08-930-264-17	Sequence 17, Appl	180	16	0.5	1247	1	US-08-451-953A-32	Sequence 32, Appl
108	16	0.5	390	4	US-08-930-264-19	Sequence 19, Appl	181	16	0.5	1247	2	US-08-445-468A-32	Sequence 32, Appl
109	16	0.5	390	4	US-08-930-264-21	Sequence 21, Appl	182	16	0.5	1247	2	US-08-461-397A-32	Sequence 32, Appl
110	16	0.5	390	4	US-08-930-264-23	Sequence 23, Appl	183	16	0.5	1247	2	US-08-912-088-32	Sequence 32, Appl
111	16	0.5	390	4	US-08-930-264-25	Sequence 25, Appl	184	16	0.5	1247	3	US-08-278-730A-32	Sequence 32, Appl
112	16	0.5	416	4	US-09-370-838-249	Sequence 249, App	185	16	0.5	1247	3	US-08-445-467-32	Sequence 32, Appl
113	16	0.5	440	4	US-09-397-787-328	Sequence 328, App	186	16	0.5	1247	3	US-08-480-515A-32	Sequence 32, Appl
c 114	16	0.5	451	4	US-09-357-251-3	Sequence 3, Appli	187	16	0.5	1247	4	US-09-170-936-32	Sequence 32, Appl
115	16	0.5	457	2	US-08-332-766A-7	Sequence 7, Appli	188	16	0.5	1247	4	US-08-461-113-32	Sequence 32, Appl
116	16	0.5	491	1	US-07-945-288-7	Sequence 7, Appli	189	16	0.5	1247	5	PCT-US93-07190-32	Sequence 32, Appl
117	16	0.5	491	1	US-08-462-831-7	Sequence 7, Appli	190	16	0.5	1247	5	PCT-US93-07231-32	Sequence 32, Appl
118	16	0.5	491	1	US-08-461-809-7	Sequence 7, Appli	191	16	0.5	1247	5	PCT-US93-08742-32	Sequence 32, Appl
119	16	0.5	491	1	US-08-461-441-7	Sequence 7, Appli	192	16	0.5	1247	5	PCT-US93-08808-32	Sequence 32, Appl
120	16	0.5	491	1	US-08-482-142-7	Sequence 7, Appli	193	16	0.5	1247	5	PCT-US93-08885-32	Sequence 32, Appl
121	16	0.5	491	2	US-08-478-572-7	Sequence 7, Appli	194	16	0.5	1254	4	US-09-322-478-25	Sequence 25, Appl
122	16	0.5	491	2	US-08-484-296-7	Sequence 7, Appli	195	16	0.5	1263	1	US-08-532-828B-11	Sequence 11, Appl
c 123	16	0.5	491	4	US-09-643-597-307	Sequence 307, App	196	16	0.5	1263	1	US-08-532-828B-12	Sequence 12, Appl
124	16	0.5	502	4	US-08-893-08518-7	Sequence 7, Appli	c 197	16	0.5	1294	4	US-09-312-038-4	Sequence 4, Appli
c 125	16	0.5	511	4	US-08-936-165A-182	Sequence 7, Appli	c 198	16	0.5	1335	4	US-09-859-698-1	Sequence 1, Appli
126	16	0.5	511	4	US-08-288-888-3	Sequence 182, App	c 199	16	0.5	1335	4	US-09-428-979-1	Sequence 1, Appli
127	16	0.5	513	1	US-08-910-075-3	Sequence 3, Appli	200	16	0.5	1345	4	US-09-372-339-1	Sequence 1, Appli
128	16	0.5	513	2	US-08-905-801A-3	Sequence 3, Appli	201	16	0.5	1345	4	US-09-372-339-2	Sequence 2, Appli
129	16	0.5	513	2	US-08-532-828B-13	Sequence 13, Appli	202	16	0.5	1345	4	US-09-144-367-3	Sequence 3, Appli
130	16	0.5	516	1	US-08-532-828B-14	Sequence 14, Appli	c 203	16	0.5	1372	6	5219569-1	Patent No. 5219569
131	16	0.5	588	4	US-09-565-808-1	Sequence 1, Appli	204	16	0.5	1393	2	US-08-765-875-1	Sequence 1, Appli
132	16	0.5	663	2	US-08-943-087-49	Sequence 49, Appl	205	16	0.5	1393	3	US-08-795-671-1	Sequence 1, Appli
133	16	0.5	663	2	US-08-943-087-51	Sequence 51, Appl	c 206	16	0.5	1401	2	US-09-134-001C-595	Sequence 595, App
134	16	0.5	663	2	US-08-943-087-53	Sequence 53, Appl	207	16	0.5	1402	2	US-09-205-860-1	Sequence 1, Appli
135	16	0.5	663	2	US-08-943-087-55	Sequence 55, Appl	208	16	0.5	1425	1	US-08-464-148-1	Sequence 1, Appli
136	16	0.5	663	2	US-08-943-087-57	Sequence 57, App	209	16	0.5	1425	1	US-08-385-500-1	Sequence 1, Appli
137	16	0.5	668	4	US-09-385-982-157	Sequence 157, App	210	16	0.5	1434	1	US-08-846-784-1	Sequence 1, Appli
138	16	0.5	700	4	US-09-236-097-7	Sequence 7, Appli	211	16	0.5	1434	4	US-09-480-921B-9	Sequence 9, Appli
c 140	16	0.5	704	3	US-09-009-156-1	Sequence 1, Appli	212	16	0.5	1475	4	US-09-643-597-122	Sequence 28, Appl
c 141	16	0.5	704	4	US-09-372-154-1	Sequence 1, Appli	c 213	16	0.5	1477	3	US-08-096-181A-7	Sequence 122, App
142	16	0.5	718	4	US-09-347-803-17	Sequence 17, Appl	214	16	0.5	1477	5	PCT-US94-08326-7	Sequence 7, Appli
143	16	0.5	745	1	US-08-036-555B-163	Sequence 163, App	215	16	0.5	1614	2	US-08-867-941-10	Sequence 7, Appli
144	16	0.5	745	1	US-08-469-569-163	Sequence 163, App	216	16	0.5	1614	2	US-08-867-941-10	Sequence 10, Appl
145	16	0.5	745	1	US-08-249-322A-163	Sequence 163, App	217	16	0.5	1623	4	US-09-074-658-10	Sequence 10, Appl
146	16	0.5	745	1	US-08-469-526A-163	Sequence 163, App	218	16	0.5	1623	2	US-08-867-941-5	Sequence 5, Appli
147	16	0.5	745	2	US-08-734-591A-163	Sequence 163, App	219	16	0.5	1623	4	US-09-074-658-5	Sequence 5, Appli
148	16	0.5	745	2	US-08-469-660-163	Sequence 163, App	220	16	0.5	1643	1	US-08-532-828B-1	Sequence 1, Appli
149	16	0.5	745	3	US-08-341-018-51	Sequence 51, Appl	221	16	0.5	1643	1	US-08-532-828B-2	Sequence 2, Appli
150	16	0.5	745	3	US-08-470-335-163	Sequence 163, App	222	16	0.5	1643	1	US-08-532-828B-8	Sequence 7, Appli
151	16	0.5	745	4	US-08-735-021-163	Sequence 163, App	223	16	0.5	1643	1	US-08-532-828B-9	Sequence 8, Appli
152	16	0.5	745	4	US-08-734-664A-163	Sequence 163, App	224	16	0.5	1643	1	US-08-532-828B-10	Sequence 9, Appli
153	16	0.5	745	4	US-08-470-339-163	Sequence 163, App	225	16	0.5	1643	1	US-08-532-828B-7	Sequence 10, Appl
154	16	0.5	745	4	US-08-467-602-163	Sequence 163, App	226	16	0.5	1643	1	US-08-700-359-8	Sequence 7, Appli
155	16	0.5	745	5	PCT-US94-05083C-159	Sequence 159, App	227	16	0.5	1643	1	US-08-700-359-8	Sequence 8, Appli
156	16	0.5	745	5	PCT-US95-06846A-163	Sequence 163, App	228	16	0.5	1643	1	US-08-700-359-10	Sequence 10, Appl
c 157	16	0.5	780	4	US-09-134-001C-1631	Sequence 1631, App	229	16	0.5	1643	1	US-08-674-168-21	Sequence 21, Appl
158	16	0.5	780	3	US-08-822-264-2	Sequence 2, Appli	230	16	0.5	1643	2	US-08-596-366-5	Sequence 5, Appli
c 159	16	0.5	874	4	US-08-881-450A-24	Sequence 24, Appl	231	16	0.5	1643	2	US-08-596-366-7	Sequence 7, Appli
c 160	16	0.5	887	4	US-09-329-350-36	Sequence 36, Appl	232	16	0.5	1643	2	US-08-967-104-5	Sequence 5, Appli
c 161	16	0.5	1044	1	US-07-975-526-3	Sequence 3, Appli	233	16	0.5	1643	2	US-08-967-104-7	Sequence 7, Appli
c 162	16	0.5	1044	4	US-07-974-409C-425	Sequence 425, App	234	16	0.5	1643	3	US-08-985-908-3	Sequence 3, Appli
163	16	0.5	1059	4	US-09-576-160B-9	Sequence 9, Appli	235	16	0.5	1643	3	US-08-985-908-4	Sequence 4, Appli
164	16	0.5	1062	4	US-09-576-160B-12	Sequence 12, Appl	236	16	0.5	1643	3	US-08-985-908-6	Sequence 6, Appli
c 165	16	0.5	1104	4	US-09-134-001C-1331	Sequence 1331, App	237	16	0.5	1643	3	US-08-852-730-13	Sequence 13, Appl
166	16	0.5	1197	4	US-09-491-362-6	Sequence 6, Appli	238	16	0.5	1643	3	US-08-852-730-12	Sequence 12, Appl
167	16	0.5	1197	4	US-09-874-562-6	Sequence 6, Appli	239	16	0.5	1643	3	US-08-852-730-15	Sequence 15, Appl
168	16	0.5	1209	4	US-08-791-115B-24	Sequence 24, Appl	240	16	0.5	1643	4	US-08-985-916-3	Sequence 3, Appli
169	16	0.5	1222	4	US-09-333-611-13	Sequence 13, Appl	241	16	0.5	1643	4	US-08-985-916-4	Sequence 4, Appli
c 170	16	0.5	1233	1	US-08-254-922-1	Sequence 1, Appli	242	16	0.5	1643	4	US-08-985-916-6	Sequence 6, Appli
c 171	16	0.5	1233	1	US-08-286-748B-1	Sequence 1, Appli	243	16	0.5	1659	2	US-08-943-087-13	Sequence 13, Appl
c 172	16	0.5	1236	1	US-07-957-039A-7	Sequence 7, Appli	244	16	0.5	1659	2	US-08-943-087-15	Sequence 15, Appl
c 173	16	0.5	1236	1	US-08-153-799-17	Sequence 17, Appli	245	16	0.5	1659	2	US-08-943-087-25	Sequence 25, Appl
							246	16	0.5	1659	2	US-08-943-087-27	Sequence 27, Appl

247	16	0.5	1659	2	US-08-943-087-29	Sequence 29, Appl	c 320	16	0.5	2301	6	5188829-2	Patent No. 5188829
248	16	0.5	1659	2	US-08-943-087-31	Sequence 31, Appl	c 321	16	0.5	2323	4	US-09-149-476-24	Sequence 24, Appl
249	16	0.5	1659	2	US-08-943-087-33	Sequence 33, Appl	c 322	16	0.5	2335	2	US-08-300-584-3	Sequence 3, Appl
250	16	0.5	1659	2	US-08-943-087-35	Sequence 35, Appl	c 323	16	0.5	2335	4	US-08-476-123-3	Sequence 3, Appl
251	16	0.5	1659	2	US-08-943-087-37	Sequence 37, Appl	c 324	16	0.5	2360	1	US-08-039-364-1	Sequence 1, Appl
252	16	0.5	1659	2	US-08-943-087-39	Sequence 39, Appl	c 325	16	0.5	2360	4	US-09-158-710-1	Sequence 1, Appl
253	16	0.5	1659	2	US-08-943-087-41	Sequence 41, Appl	c 326	16	0.5	2450	1	US-09-491-522-2	Sequence 2, Appl
254	16	0.5	1659	2	US-08-943-087-43	Sequence 43, Appl	c 327	16	0.5	2456	1	US-07-882-711-1	Sequence 1, Appl
255	16	0.5	1659	2	US-08-943-087-45	Sequence 45, Appl	c 328	16	0.5	2456	2	US-08-462-174-1	Sequence 1, Appl
256	16	0.5	1674	2	US-08-999-811-1	Sequence 1, Appl	c 329	16	0.5	2457	2	US-08-723-415B-12	Sequence 12, Appl
257	16	0.5	1674	3	US-09-042-105-1	Sequence 1, Appl	c 330	16	0.5	2457	3	US-09-189-627A-12	Sequence 12, Appl
258	16	0.5	1674	4	US-09-134-001C-1555	Sequence 1555, Ap	c 331	16	0.5	2457	4	US-09-710-861-12	Sequence 12, Appl
259	16	0.5	1674	5	PCT-US96-09001-1	Sequence 1, Appl	c 332	16	0.5	2457	1	US-08-139-937-13	Sequence 13, Appl
260	16	0.5	1702	1	US-07-616-022C-1	Sequence 1, Appl	c 333	16	0.5	2492	5	PCT-US93-11310-13	Sequence 13, Appl
261	16	0.5	1732	4	US-09-449-335-1	Sequence 1, Appl	c 334	16	0.5	2517	1	US-08-306-691B-18	Sequence 18, Appl
262	16	0.5	1732	4	US-09-449-335-5	Sequence 5, Appl	c 335	16	0.5	2517	1	US-08-385-142-2	Sequence 2, Appl
263	16	0.5	1809	1	US-08-455-001-1	Sequence 1, Appl	c 336	16	0.5	2517	3	US-08-481-814A-1	Sequence 1, Appl
264	16	0.5	1809	5	PCT-US95-11869-1	Sequence 1, Appl	c 337	16	0.5	2517	3	US-08-801-092-2	Sequence 2, Appl
265	16	0.5	1837	2	US-08-750-134A-4	Sequence 4, Appl	c 338	16	0.5	2517	4	US-09-517-584A-3	Sequence 3, Appl
266	16	0.5	1837	4	US-09-363-745-4	Sequence 4, Appl	c 339	16	0.5	2517	4	US-09-315-113-2	Sequence 2, Appl
267	16	0.5	1868	2	US-08-960-022-5	Sequence 5, Appl	c 340	16	0.5	2572	4	US-09-221-017B-225	Sequence 225, App
268	16	0.5	1882	1	US-08-696-349-1	Sequence 1, Appl	c 341	16	0.5	2671	2	US-08-408-519-1	Sequence 1, Appl
269	16	0.5	1882	5	PCT-US96-13156-1	Sequence 1, Appl	c 342	16	0.5	2671	5	PCT-US95-03552-1	Sequence 1, Appl
270	16	0.5	1886	1	US-07-980-526-1	Sequence 1, Appl	c 343	16	0.5	2730	2	US-08-811-897A-39	Sequence 39, Appl
271	16	0.5	1894	1	US-08-476-008-66	Sequence 66, Appl	c 344	16	0.5	2730	2	US-08-855-213-39	Sequence 39, Appl
272	16	0.5	1894	1	US-08-306-063-66	Sequence 66, Appl	c 345	16	0.5	2730	4	US-09-201-474-39	Sequence 39, Appl
273	16	0.5	1894	1	US-08-833-485-66	Sequence 66, Appl	c 346	16	0.5	2745	1	US-08-385-370-3	Sequence 3, Appl
274	16	0.5	1894	4	US-09-137-440-66	Sequence 66, Appl	c 347	16	0.5	2814	2	US-08-811-897A-38	Sequence 38, Appl
275	16	0.5	1908	2	US-08-909-965C-17	Sequence 17, Appl	c 348	16	0.5	2814	2	US-08-855-213-38	Sequence 38, Appl
276	16	0.5	1929	4	US-09-040-229B-7	Sequence 7, Appl	c 349	16	0.5	2814	4	US-09-201-474-38	Sequence 38, Appl
277	16	0.5	1953	1	US-08-315-468-5	Sequence 5, Appl	c 350	16	0.5	2814	4	US-09-177-650-90	Sequence 90, Appl
278	16	0.5	1953	4	US-07-941-650A-3	Sequence 3, Appl	c 351	16	0.5	2832	3	US-08-819-177-2	Sequence 2, Appl
279	16	0.5	1988	2	US-08-712-948-8	Sequence 8, Appl	c 352	16	0.5	2838	4	US-08-246-489-1	Sequence 1, Appl
280	16	0.5	1996	1	US-08-385-370-1	Sequence 1, Appl	c 353	16	0.5	2850	2	US-08-224-482-7	Sequence 7, Appl
281	16	0.5	1997	3	US-08-795-430-7	Sequence 7, Appl	c 354	16	0.5	2935	4	US-09-480-921B-27	Sequence 27, Appl
282	16	0.5	1997	4	US-08-510-133A-34	Sequence 34, Appl	c 355	16	0.5	2953	2	US-08-859-201-1	Sequence 1, Appl
283	16	0.5	1997	4	US-09-355-700-7	Sequence 7, Appl	c 356	16	0.5	2953	3	US-08-906-360-2	Sequence 2, Appl
284	16	0.5	2003	1	US-08-601-132-32	Sequence 32, Appl	c 357	16	0.5	3123	1	US-08-094-889-2	Sequence 2, Appl
285	16	0.5	2003	1	US-08-036-555B-21	Sequence 21, Appl	c 358	16	0.5	3200	1	US-08-444-405-1	Sequence 1, Appl
286	16	0.5	2003	1	US-08-469-569-21	Sequence 21, Appl	c 359	16	0.5	3200	1	US-08-384-850-1	Sequence 1, Appl
287	16	0.5	2003	1	US-08-249-322A-21	Sequence 21, Appl	c 360	16	0.5	3278	1	US-08-484-105-13	Sequence 13, Appl
288	16	0.5	2003	1	US-08-469-526A-21	Sequence 21, Appl	c 361	16	0.5	3278	1	US-08-484-106-13	Sequence 13, Appl
289	16	0.5	2003	2	US-08-734-591A-21	Sequence 21, Appl	c 362	16	0.5	3358	2	US-08-469-537A-104	Sequence 104, App
290	16	0.5	2003	2	US-08-469-660-21	Sequence 21, Appl	c 363	16	0.5	3516	2	US-08-943-087-1	Sequence 1, Appl
291	16	0.5	2003	3	US-08-341-018-71	Sequence 71, Appl	c 364	16	0.5	3564	4	US-09-347-878-15	Sequence 15, Appl
292	16	0.5	2003	3	US-08-470-335-21	Sequence 21, Appl	c 365	16	0.5	3566	2	US-08-415-788-4	Sequence 4, Appl
293	16	0.5	2003	4	US-08-735-021-21	Sequence 21, Appl	c 366	16	0.5	3566	2	US-08-415-788-2	Sequence 2, Appl
294	16	0.5	2003	4	US-08-734-664A-21	Sequence 21, Appl	c 367	16	0.5	3567	2	US-08-415-788-6	Sequence 6, Appl
295	16	0.5	2003	4	US-08-470-339-21	Sequence 21, Appl	c 368	16	0.5	3567	2	US-08-415-788-2	Sequence 2, Appl
296	16	0.5	2003	4	US-08-467-602-21	Sequence 21, Appl	c 369	16	0.5	3567	4	US-08-104-072B-6	Sequence 6, Appl
297	16	0.5	2003	5	PCT-US94-05083C-21	Sequence 21, Appl	c 370	16	0.5	3627	1	US-08-351-413-7	Sequence 7, Appl
298	16	0.5	2003	5	PCT-US95-06846A-21	Sequence 21, Appl	c 371	16	0.5	3627	2	US-09-025-583-7	Sequence 7, Appl
299	16	0.5	2007	4	US-09-149-476-135	Sequence 135, App	c 372	16	0.5	3627	2	US-08-737-524B-26	Sequence 26, Appl
300	16	0.5	2009	4	US-09-446-504-17	Sequence 17, Appl	c 373	16	0.5	3639	2	US-08-737-524B-26	Sequence 26, Appl
301	16	0.5	2009	4	US-09-712-266-17	Sequence 17, Appl	c 374	16	0.5	3639	2	US-08-737-524B-26	Sequence 26, Appl
302	16	0.5	2023	4	US-09-491-522-6	Sequence 6, Appl	c 375	16	0.5	3651	4	US-07-852-132A-13	Sequence 13, Appl
303	16	0.5	2031	4	US-08-706-054A-1	Sequence 1, Appl	c 376	16	0.5	3666	6	5248670-3	Patent No. 5248670
304	16	0.5	2031	4	US-08-706-054A-1	Sequence 1, Appl	c 377	16	0.5	3765	3	US-07-705-490-1	Sequence 1, Appl
305	16	0.5	2077	1	US-08-217-327-7	Sequence 7, Appl	c 378	16	0.5	3765	4	US-07-751-891B-1	Sequence 1, Appl
306	16	0.5	2117	1	US-08-431-080-23	Sequence 23, Appl	c 379	16	0.5	3989	1	US-08-327-494A-1	Sequence 1, Appl
307	16	0.5	2117	2	US-08-938-53A-23	Sequence 23, Appl	c 380	16	0.5	3989	1	US-08-327-494A-3	Sequence 3, Appl
308	16	0.5	2117	4	US-09-345-29A-23	Sequence 23, Appl	c 381	16	0.5	3989	5	PCT-US95-13659-1	Sequence 1, Appl
309	16	0.5	2156	2	US-08-959-011-2	Sequence 2, Appl	c 382	16	0.5	3989	5	PCT-US95-13659-3	Sequence 3, Appl
310	16	0.5	2180	2	US-08-755-559-2	Sequence 2, Appl	c 383	16	0.5	4010	2	US-08-785-310A-3	Sequence 3, Appl
311	16	0.5	2180	3	US-09-210-47A-2	Sequence 2, Appl	c 384	16	0.5	4098	4	US-09-268-866-1	Sequence 1, Appl
312	16	0.5	2180	4	US-09-539-77A-2	Sequence 2, Appl	c 385	16	0.5	4134	4	US-09-162-021B-1	Sequence 1, Appl
313	16	0.5	2190	4	US-08-893-654B-1	Sequence 1, Appl	c 386	16	0.5	4362	2	US-09-687-477-17	Sequence 17, Appl
314	16	0.5	2263	1	US-08-176-126B-1	Sequence 1, Appl	c 387	16	0.5	4362	2	US-08-455-073A-1	Sequence 1, Appl
315	16	0.5	2263	2	US-08-669-435-1	Sequence 1, Appl	c 388	16	0.5	4402	4	US-09-484-970B-135	Sequence 135, App
316	16	0.5	2263	5	PCT-US94-14431A-1	Sequence 1, Appl	c 389	16	0.5	4804	4	US-09-066-047-6	Sequence 6, Appl
317	16	0.5	2268	4	US-08-675-773B-4	Sequence 4, Appl	c 390	16	0.5	4935	2	US-08-631-097-3	Sequence 3, Appl
318	16	0.5	2289	4	US-09-312-038-3	Sequence 3, Appl	c 391	16	0.5	5288	2	US-08-540-406-18	Sequence 18, Appl
319	16	0.5	2294	4	US-09-643-597-123	Sequence 123, App	c 392	16	0.5	5288	3	US-08-656-055-18	Sequence 18, Appl

393	16	0.5	5288	4	US-08-954-668-18	Sequence 18, Appl	466	15	0.5	25	2	US-08-775-607-3	Sequence 3, Appl
394	16	0.5	5288	4	US-08-918-658-18	Sequence 18, Appl	467	15	0.5	25	3	US-08-622-277A-17	Sequence 17, Appl
395	16	0.5	5288	5	PCT-US95-13233-18	Sequence 18, Appl	468	15	0.5	25	5	PCT-US93-06828-3	Sequence 3, Appl
396	16	0.5	5886	4	US-08-810-712-9	Sequence 9, Appl	469	15	0.5	28	3	US-08-863-813A-38	Sequence 38, Appl
397	16	0.5	6211	4	US-08-961-527-8	Sequence 8, Appl	470	15	0.5	28	4	US-09-025-580-6	Sequence 6, Appl
c 398	16	0.5	6370	4	US-09-306-595C-2	Sequence 2, Appl	c 471	15	0.5	28	4	US-08-676-318A-38	Sequence 38, Appl
c 399	16	0.5	6692	4	US-09-491-522-1	Sequence 6, Appl	c 472	15	0.5	30	1	US-08-246-978A-2	Sequence 2, Appl
400	16	0.5	6714	4	US-09-299-141-6	Sequence 1, Appl	c 473	15	0.5	30	1	US-08-068-747-2	Sequence 2, Appl
401	16	0.5	6803	3	US-08-665-259-19	Sequence 19, Appl	474	15	0.5	33	1	US-08-440-814A-2	Sequence 2, Appl
402	16	0.5	6803	3	US-08-762-500-19	Sequence 19, Appl	475	15	0.5	33	1	US-08-068-747-7	Sequence 7, Appl
403	16	0.5	6924	4	US-09-299-141-9	Sequence 9, Appl	476	15	0.5	38	3	US-08-863-813A-51	Sequence 51, Appl
404	16	0.5	6924	4	US-09-299-141-10	Sequence 10, Appl	c 477	15	0.5	41	4	US-09-402-631A-1	Sequence 1, Appl
405	16	0.5	6924	4	US-09-299-141-11	Sequence 11, Appl	c 478	15	0.5	45	4	US-09-497-933A-22	Sequence 22, Appl
c 406	16	0.5	7176	4	US-09-221-017B-626	Sequence 626, App	c 479	15	0.5	48	4	US-09-497-933A-19	Sequence 19, Appl
407	16	0.5	7336	4	US-09-306-446C-1	Sequence 1, Appl	480	15	0.5	50	4	US-09-930-181-13	Sequence 13, Appl
408	16	0.5	7641	2	US-08-867-941-6	Sequence 6, Appl	481	15	0.5	50	4	US-09-930-181-14	Sequence 14, Appl
409	16	0.5	7641	4	US-09-074-658-6	Sequence 6, Appl	482	15	0.5	60	4	US-09-497-933A-21	Sequence 21, Appl
410	16	0.5	7650	2	US-08-867-941-1	Sequence 1, Appl	c 483	15	0.5	63	4	US-09-497-933A-18	Sequence 18, Appl
411	16	0.5	7650	4	US-09-074-658-1	Sequence 1, Appl	c 484	15	0.5	63	4	US-09-497-933A-23	Sequence 23, Appl
c 412	16	0.5	8040	1	US-08-596-291-1	Sequence 1, Appl	485	15	0.5	63	4	US-09-497-933A-25	Sequence 25, Appl
c 413	16	0.5	8040	1	US-09-100-804-1	Sequence 1, Appl	486	15	0.5	65	3	US-08-795-430-17	Sequence 17, Appl
c 414	16	0.5	8043	5	PCT-US94-09943-1	Sequence 5, Appl	487	15	0.5	65	4	US-09-355-700-17	Sequence 17, Appl
c 415	16	0.5	8050	4	US-09-491-362-11	Sequence 11, Appl	488	15	0.5	66	3	US-08-516-859A-56	Sequence 56, Appl
c 416	16	0.5	8050	4	US-09-491-362-11	Sequence 11, Appl	489	15	0.5	66	4	US-09-586-472-96	Sequence 96, Appl
c 417	16	0.5	8119	4	US-09-290-640-45	Sequence 45, Appl	490	15	0.5	66	4	US-09-528-706-96	Sequence 96, Appl
418	16	0.5	8906	2	US-08-826-267-1	Sequence 1, Appl	c 491	15	0.5	69	3	US-08-480-173A-31	Sequence 31, Appl
419	16	0.5	8982	3	US-08-976-255-5	Sequence 5, Appl	492	15	0.5	69	3	US-08-480-173A-32	Sequence 32, Appl
420	16	0.5	9837	4	US-09-221-017B-636	Sequence 636, App	c 493	15	0.5	69	3	US-08-484-408A-31	Sequence 31, Appl
c 421	16	0.5	9972	3	US-08-836-022A-3	Sequence 3, Appl	c 494	15	0.5	77	2	US-08-484-408A-32	Sequence 32, Appl
c 422	16	0.5	9972	3	US-09-427-048A-3	Sequence 3, Appl	c 495	15	0.5	77	2	US-08-180-524-5	Sequence 5, Appl
423	16	0.5	10299	2	US-08-477-451-1	Sequence 1, Appl	c 496	15	0.5	77	2	US-08-975-166-5	Sequence 5, Appl
c 424	16	0.5	10299	2	US-08-477-451-5	Sequence 5, Appl	c 497	15	0.5	78	4	US-09-497-933A-24	Sequence 24, Appl
c 425	16	0.5	10348	2	US-08-457-273B-41	Sequence 41, Appl	498	15	0.5	81	4	US-09-497-933A-20	Sequence 20, Appl
c 426	16	0.5	10348	3	US-08-556-419-13	Sequence 13, Appl	499	15	0.5	84	3	US-08-795-430-19	Sequence 19, Appl
c 427	16	0.5	10348	4	US-09-041-886-14	Sequence 14, Appl	500	15	0.5	84	4	US-09-355-700-19	Sequence 19, Appl
c 428	16	0.5	10366	1	US-08-246-982A-5	Sequence 5, Appl	501	15	0.5	90	1	US-08-398-617-15	Sequence 15, Appl
c 429	16	0.5	10366	1	US-08-453-265-5	Sequence 5, Appl	502	15	0.5	90	2	US-08-398-615-15	Sequence 15, Appl
430	16	0.5	11580	4	US-09-334-220-4	Sequence 4, Appl	503	15	0.5	90	4	US-07-827-691A-4	Sequence 4, Appl
431	16	0.5	11932	2	US-08-477-451-25	Sequence 25, Appl	504	15	0.5	90	4	US-08-397-303-15	Sequence 15, Appl
432	16	0.5	35408	4	US-08-973-334-3	Sequence 3, Appl	505	15	0.5	96	4	US-09-025-769B-291	Sequence 291, App
433	16	0.5	35408	4	US-09-563-869A-3	Sequence 3, Appl	c 506	15	0.5	103	3	US-08-480-173A-12	Sequence 12, Appl
434	16	0.5	35408	4	US-08-549-489-3	Sequence 3, Appl	c 507	15	0.5	103	3	US-08-484-408A-12	Sequence 12, Appl
435	16	0.5	35524	3	US-08-923-137-1	Sequence 1, Appl	c 508	15	0.5	146	2	US-08-180-524-3	Sequence 3, Appl
436	16	0.5	72604	4	US-09-268-992-7	Sequence 7, Appl	c 509	15	0.5	146	2	US-08-975-166-3	Sequence 3, Appl
437	16	0.5	72604	4	US-09-657-474-7	Sequence 7, Appl	c 510	15	0.5	161	4	US-09-404-879A-35	Sequence 35, Appl
438	16	0.5	87350	3	US-08-781-891-79	Sequence 79, Appl	c 511	15	0.5	168	4	US-09-328-111-222	Sequence 222, App
439	16	0.5	87543	4	US-09-791-211-3	Sequence 3, Appl	c 512	15	0.5	180	1	US-08-343-281A-20	Sequence 20, Appl
440	15	0.5	15	2	US-08-863-639A-21	Sequence 21, Appl	c 513	15	0.5	186	3	US-09-188-930-95	Sequence 95, Appl
c 441	15	0.5	17	3	US-08-909-742-3	Sequence 3, Appl	514	15	0.5	186	4	US-09-402-328-4	Sequence 4, Appl
c 442	15	0.5	17	3	US-08-909-742-4	Sequence 4, Appl	515	15	0.5	189	4	US-09-568-816A-3	Sequence 3, Appl
c 443	15	0.5	17	4	US-09-412-289-3	Sequence 3, Appl	516	15	0.5	192	4	US-09-134-001C-49	Sequence 49, Appl
c 444	15	0.5	17	4	US-09-412-289-4	Sequence 4, Appl	517	15	0.5	195	1	US-08-324-243-20	Sequence 20, Appl
445	15	0.5	18	2	US-08-857-946-14	Sequence 14, Appl	518	15	0.5	195	1	US-08-532-390-20	Sequence 20, Appl
446	15	0.5	18	3	US-08-970-740-14	Sequence 14, Appl	519	15	0.5	195	3	US-08-717-294-20	Sequence 20, Appl
447	15	0.5	18	3	US-09-143-212-45	Sequence 45, Appl	520	15	0.5	195	5	PCT-US95-11511-20	Sequence 20, Appl
448	15	0.5	20	2	US-08-465-485A-28	Sequence 28, Appl	c 521	15	0.5	200	5	US-08-983-502-3	Sequence 3, Appl
449	15	0.5	20	3	US-09-080-285-28	Sequence 28, Appl	c 522	15	0.5	200	5	PCT-US96-10521-3	Sequence 3, Appl
450	15	0.5	20	4	US-09-030-701-65	Sequence 65, Appl	c 523	15	0.5	249	1	US-07-872-644-40	Sequence 40, Appl
451	15	0.5	20	4	US-09-593-711A-37	Sequence 37, Appl	c 524	15	0.5	249	1	US-08-297-494-40	Sequence 40, Appl
452	15	0.5	20	4	US-09-082-649B-57	Sequence 57, Appl	c 525	15	0.5	249	1	US-08-297-510-40	Sequence 40, Appl
453	15	0.5	20	4	US-09-724-426-28	Sequence 28, Appl	c 526	15	0.5	249	1	US-08-479-532-40	Sequence 40, Appl
c 454	15	0.5	21	2	US-08-863-639A-52	Sequence 52, Appl	c 527	15	0.5	249	1	US-08-455-525-40	Sequence 40, Appl
c 455	15	0.5	21	2	US-08-863-639A-55	Sequence 55, Appl	c 528	15	0.5	249	1	US-08-455-525-40	Sequence 40, Appl
456	15	0.5	21	2	US-08-863-639A-56	Sequence 56, Appl	c 529	15	0.5	249	5	US-09-139-491-40	Sequence 40, Appl
c 457	15	0.5	21	2	US-08-863-639A-67	Sequence 67, Appl	c 530	15	0.5	249	5	PCT-US92-03222-40	Sequence 40, Appl
458	15	0.5	21	2	US-08-863-639A-68	Sequence 68, Appl	c 531	15	0.5	256	4	US-08-483-533-11	Sequence 11, Appl
459	15	0.5	21	2	US-08-863-639A-71	Sequence 71, Appl	c 532	15	0.5	256	4	US-09-283-471A-11	Sequence 11, Appl
460	15	0.5	21	2	US-08-416-214A-11	Sequence 11, Appl	533	15	0.5	261	4	US-09-009-816-7	Sequence 7, Appl
c 461	15	0.5	24	2	US-08-570-155-16	Sequence 16, Appl	534	15	0.5	288	2	US-08-716-942-16	Sequence 16, Appl
462	15	0.5	24	2	US-08-570-155-17	Sequence 17, Appl	535	15	0.5	288	4	US-09-130-337A-16	Sequence 16, Appl
463	15	0.5	25	1	US-08-374-144-3	Sequence 3, Appl	536	15	0.5	318	2	US-08-646-981-4	Sequence 4, Appl
464	15	0.5	25	1	US-08-775-164-3	Sequence 3, Appl	537	15	0.5	329	1	US-08-148-910-2	Sequence 2, Appl
465	15	0.5	25	2	US-08-775-609-3	Sequence 3, Appl	538	15	0.5	329	1	US-08-148-910-13	Sequence 13, Appl



539	15	0.5	329	1	US-08-448-937A-2	Sequence 2, Appl	612	15	0.5	800	3	US-09-027-449-69	Sequence 69, Appl
540	15	0.5	329	1	US-08-448-937A-13	Sequence 13, Appl	613	15	0.5	800	3	US-09-026-985-69	Sequence 69, Appl
c 541	15	0.5	330	1	US-09-325-932A-93	Sequence 93, Appl	614	15	0.5	800	3	US-09-121-952A-69	Sequence 69, Appl
542	15	0.5	346	2	US-08-761-277A-50	Sequence 50, Appl	615	15	0.5	800	4	US-09-234-340A-69	Sequence 69, Appl
543	15	0.5	349	4	US-09-198-119C-94	Sequence 94, Appl	616	15	0.5	804	2	US-08-835-099A-10	Sequence 10, Appl
544	15	0.5	364	4	US-09-221-017B-481	Sequence 481, App	617	15	0.5	804	3	US-09-157-349-10	Sequence 10, Appl
c 545	15	0.5	367	4	US-09-325-932A-2	Sequence 2, Appl	c 618	15	0.5	819	3	US-08-792-014-2	Sequence 2, Appl
c 546	15	0.5	423	1	US-08-470-179-190	Sequence 190, App	c 619	15	0.5	819	4	US-09-443-948-2	Sequence 2, Appl
c 547	15	0.5	425	1	US-08-664-596B-12	Sequence 12, Appl	620	15	0.5	826	4	US-08-998-416-364	Sequence 364, App
c 548	15	0.5	425	4	US-09-397-787-249	Sequence 249, App	621	15	0.5	829	3	US-08-961-083-133	Sequence 133, App
c 549	15	0.5	431	4	US-08-483-533-17	Sequence 17, Appl	622	15	0.5	832	2	US-08-686-417-1	Sequence 1, Appl
c 550	15	0.5	431	4	US-09-283-471A-17	Sequence 17, Appl	623	15	0.5	834	4	US-09-134-001C-133	Sequence 133, App
551	15	0.5	450	4	US-09-370-838-145	Sequence 145, App	624	15	0.5	836	3	US-08-674-984-1	Sequence 1, Appl
c 552	15	0.5	457	3	US-08-952-664-10	Sequence 10, Appl	625	15	0.5	836	3	US-08-674-984-2	Sequence 2, Appl
c 553	15	0.5	457	4	US-09-487-874-10	Sequence 10, Appl	626	15	0.5	836	5	PCT-US95-15601-1	Sequence 1, Appl
c 554	15	0.5	463	4	US-09-221-017B-201	Sequence 201, App	627	15	0.5	836	5	PCT-US95-15601-2	Sequence 2, Appl
c 555	15	0.5	465	4	US-08-171-209-29	Sequence 29, Appl	c 628	15	0.5	856	2	US-09-057-762-23	Sequence 23, Appl
c 556	15	0.5	500	4	US-09-575-574-2	Sequence 2, Appl	c 629	15	0.5	856	3	US-08-326-119A-23	Sequence 23, Appl
c 557	15	0.5	501	4	US-09-404-879A-118	Sequence 118, App	c 630	15	0.5	859	4	US-09-177-650-117	Sequence 117, App
c 558	15	0.5	502	4	US-08-998-416-124	Sequence 124, App	c 631	15	0.5	862	4	US-08-818-112-9	Sequence 9, Appl
559	15	0.5	506	1	US-08-398-617-13	Sequence 13, Appl	c 632	15	0.5	862	4	US-08-818-111-9	Sequence 9, Appl
560	15	0.5	506	2	US-08-398-615-13	Sequence 13, Appl	c 633	15	0.5	862	4	US-09-056-556-9	Sequence 9, Appl
561	15	0.5	506	4	US-08-397-303-13	Sequence 13, Appl	c 634	15	0.5	862	4	US-09-072-596-9	Sequence 9, Appl
562	15	0.5	510	4	US-09-452-239-9	Sequence 9, Appl	635	15	0.5	869	1	US-08-249-671A-7	Sequence 7, Appl
563	15	0.5	516	2	US-08-943-814-1	Sequence 1, Appl	636	15	0.5	875	2	US-08-778-912A-5	Sequence 5, Appl
564	15	0.5	519	2	US-08-581-528A-5	Sequence 5, Appl	637	15	0.5	875	4	US-09-541-941B-5	Sequence 5, Appl
565	15	0.5	519	5	PCT-US94-07799-5	Sequence 5, Appl	638	15	0.5	879	1	US-08-249-671A-10	Sequence 10, Appl
566	15	0.5	530	3	US-08-758-662-4	Sequence 4, Appl	c 639	15	0.5	879	2	US-09-024-848-1	Sequence 1, Appl
567	15	0.5	531	4	US-09-191-608-8	Sequence 8, Appl	c 640	15	0.5	879	4	US-09-348-116A-1	Sequence 1, Appl
568	15	0.5	539	2	US-08-702-703-1	Sequence 1, Appl	c 641	15	0.5	885	3	US-08-545-196B-20	Sequence 20, Appl
c 574	15	0.5	541	2	US-08-943-814-9	Sequence 9, Appl	642	15	0.5	887	4	US-09-329-350-36	Sequence 36, Appl
c 575	15	0.5	543	1	US-08-468-036-11	Sequence 11, Appl	643	15	0.5	897	2	US-08-486-663A-19	Sequence 19, Appl
c 576	15	0.5	543	2	US-08-376-843-11	Sequence 11, Appl	644	15	0.5	897	3	US-08-767-942A-24	Sequence 24, Appl
571	15	0.5	543	2	US-09-109-266-3	Sequence 3, Appl	645	15	0.5	897	4	US-09-134-001C-645	Sequence 645, App
572	15	0.5	544	4	US-09-134-001C-3	Sequence 3, Appl	c 646	15	0.5	903	4	US-09-957-351-6	Sequence 6, Appl
573	15	0.5	564	4	US-09-449-285A-12	Sequence 12, Appl	c 647	15	0.5	906	4	US-09-453-702B-117	Sequence 117, App
c 574	15	0.5	582	1	US-08-722-001-23	Sequence 23, Appl	648	15	0.5	906	4	US-09-134-001C-78	Sequence 78, Appl
c 575	15	0.5	587	4	US-09-221-017B-783	Sequence 783, App	649	15	0.5	906	4	US-09-105-390-39	Sequence 39, Appl
c 576	15	0.5	593	1	US-08-443-568B-13	Sequence 13, Appl	c 650	15	0.5	912	4	US-09-134-001C-340	Sequence 340, App
578	15	0.5	593	5	PCT-US94-06997-13	Sequence 13, Appl	c 651	15	0.5	919	4	US-09-342-084-9	Sequence 9, Appl
c 579	15	0.5	594	2	US-09-280-116-240	Sequence 240, App	652	15	0.5	925	2	US-08-916-901-4	Sequence 4, Appl
580	15	0.5	595	2	US-09-109-266-15	Sequence 15, Appl	c 653	15	0.5	925	4	US-09-154-602-4	Sequence 4, Appl
581	15	0.5	595	4	US-09-228-986-42	Sequence 42, Appl	c 654	15	0.5	925	4	US-09-257-179-17	Sequence 17, Appl
582	15	0.5	615	4	US-08-998-416-1063	Sequence 1063, Ap	655	15	0.5	925	4	US-09-027-449-59	Sequence 59, Appl
c 583	15	0.5	633	1	US-08-234-783-1	Sequence 1, Appl	656	15	0.5	927	3	US-08-804-444A-59	Sequence 59, Appl
c 584	15	0.5	633	1	US-08-456-907-1	Sequence 1, Appl	657	15	0.5	927	3	US-09-026-985-59	Sequence 59, Appl
585	15	0.5	633	3	US-08-950-925-1	Sequence 1, Appl	658	15	0.5	927	3	US-09-121-952A-59	Sequence 59, Appl
c 586	15	0.5	633	5	PCT-US95-05523-1	Sequence 3, Appl	659	15	0.5	927	4	US-09-234-340A-59	Sequence 59, Appl
587	15	0.5	635	4	US-09-385-982-3	Sequence 16, Appl	c 661	15	0.5	930	4	US-09-105-058C-17	Sequence 17, Appl
588	15	0.5	640	2	US-08-835-099A-16	Sequence 16, Appl	c 662	15	0.5	936	4	US-09-416-509C-2	Sequence 2, Appl
589	15	0.5	640	3	US-09-157-349-16	Sequence 16, Appl	c 663	15	0.5	969	4	US-09-134-001C-1227	Sequence 1227, Ap
590	15	0.5	645	3	US-08-985-526-22	Sequence 22, Appl	664	15	0.5	970	1	US-08-148-910-3	Sequence 3, Appl
591	15	0.5	645	4	US-09-535-008-30	Sequence 30, Appl	665	15	0.5	991	4	US-08-448-937A-3	Sequence 3, Appl
592	15	0.5	651	2	US-08-361-858-1	Sequence 1, Appl	666	15	0.5	993	4	US-09-344-529-3	Sequence 3, Appl
593	15	0.5	651	2	US-08-961-858-3	Sequence 3, Appl	c 667	15	0.5	993	4	US-09-105-390-55	Sequence 55, Appl
594	15	0.5	651	3	US-09-089-593-1	Sequence 1, Appl	c 668	15	0.5	1001	4	US-09-641-638-159	Sequence 159, App
595	15	0.5	651	3	US-09-089-593-3	Sequence 3, Appl	c 669	15	0.5	1017	4	US-08-957-351-5	Sequence 5, Appl
596	15	0.5	651	3	US-08-993-380-5	Sequence 5, Appl	c 670	15	0.5	1023	1	US-08-252-966B-16	Sequence 16, Appl
597	15	0.5	666	1	US-08-663-023-16	Sequence 16, Appl	671	15	0.5	1026	4	US-07-751-891B-24	Sequence 24, Appl
c 598	15	0.5	670	4	US-09-009-816-3	Sequence 3, Appl	672	15	0.5	1027	4	US-09-344-529-1	Sequence 1, Appl
c 599	15	0.5	674	3	US-09-069-811-1	Sequence 1, Appl	673	15	0.5	1028	4	US-08-118-200-1	Sequence 1, Appl
c 600	15	0.5	674	4	US-09-543-106-1	Sequence 1, Appl	674	15	0.5	1028	4	US-08-458-745-1	Sequence 1, Appl
601	15	0.5	711	4	US-08-998-416-1143	Sequence 1143, Ap	675	15	0.5	1037	1	US-08-240-124-1	Sequence 1, Appl
c 602	15	0.5	737	1	US-08-230-574-3	Sequence 3, Appl	676	15	0.5	1037	1	US-08-453-943-1	Sequence 1, Appl
c 603	15	0.5	769	4	US-09-535-008-50	Sequence 50, Appl	677	15	0.5	1037	2	US-09-057-121-1	Sequence 1, Appl
604	15	0.5	789	4	US-09-020-956-32	Sequence 32, Appl	678	15	0.5	1037	4	US-09-358-734-1	Sequence 1, Appl
605	15	0.5	789	4	US-09-030-607-32	Sequence 32, Appl	679	15	0.5	1051	4	US-09-245-041-10	Sequence 10, Appl
606	15	0.5	789	4	US-09-605-785-32	Sequence 32, Appl	680	15	0.5	1056	4	US-09-524-162-1	Sequence 1, Appl
607	15	0.5	789	4	US-09-439-313-32	Sequence 32, Appl	c 681	15	0.5	1058	4	US-09-452-239-11	Sequence 11, Appl
608	15	0.5	789	4	US-09-352-616A-32	Sequence 32, Appl	c 682	15	0.5	1068	4	US-09-403-768-5	Sequence 5, Appl
609	15	0.5	789	4	US-09-232-149A-32	Sequence 32, Appl	c 683	15	0.5	1070	4	US-09-470-443-7	Sequence 7, Appl
610	15	0.5	794	2	US-08-675-921B-8	Sequence 8, Appl	c 684	15	0.5	1083	2	US-08-319-866-1	Sequence 1, Appl
c 611	15	0.5	800	1	US-08-230-574-1	Sequence 1, Appl							



c 831	15	0.5	1700	5	PCT-US95-04801-1	Sequence 1, Appl	904	15	0.5	2140	2	US-08-722-190-1	Sequence 1, Appl
c 832	15	0.5	1701	4	US-09-232-468A-21	Sequence 21, Appl	905	15	0.5	2140	3	US-08-244-354-1	Sequence 1, Appl
c 833	15	0.5	1727	3	US-08-999-733-2	Sequence 2, Appl	906	15	0.5	2140	3	US-09-206-899-1	Sequence 1, Appl
c 834	15	0.5	1730	4	US-09-575-574-5	Sequence 5, Appl	907	15	0.5	2140	4	US-09-444-783-1	Sequence 1, Appl
c 835	15	0.5	1739	4	US-09-594-193-1	Sequence 1, Appl	908	15	0.5	2140	4	US-09-688-415-1	Sequence 1, Appl
c 836	15	0.5	1741	1	US-08-565-655-5	Sequence 5, Appl	909	15	0.5	2140	5	PCT-US95-04203-1	Sequence 5, Appl
c 837	15	0.5	1752	4	US-09-152-060-38	Sequence 38, Appl	910	15	0.5	2143	3	US-09-097-309-5	Sequence 9, Appl
c 838	15	0.5	1762	2	US-08-742-621-2	Sequence 2, Appl	911	15	0.5	2143	3	US-09-097-171A-9	Sequence 9, Appl
c 839	15	0.5	1769	3	US-08-765-268A-1	Sequence 1, Appl	912	15	0.5	2143	4	US-09-460-587-5	Sequence 5, Appl
c 840	15	0.5	1770	3	US-09-006-636-1	Sequence 1, Appl	913	15	0.5	2156	4	US-08-965-762-16	Sequence 16, Appl
c 841	15	0.5	1770	4	US-09-006-632-1	Sequence 1, Appl	914	15	0.5	2156	4	US-09-911-927-16	Sequence 16, Appl
c 842	15	0.5	1770	4	US-09-325-274-1	Sequence 1, Appl	c 915	15	0.5	2156	4	US-09-911-927-18	Sequence 18, Appl
c 843	15	0.5	1776	1	US-08-722-001-29	Sequence 29, Appl	916	15	0.5	2156	4	US-09-911-882-16	Sequence 16, Appl
c 844	15	0.5	1778	4	US-08-934-386-4	Sequence 4, Appl	c 917	15	0.5	2156	4	US-09-911-882-18	Sequence 18, Appl
c 845	15	0.5	1791	4	US-09-134-001C-526	Sequence 526, App	918	15	0.5	2178	1	US-08-463-587A-24	Sequence 24, Appl
c 846	15	0.5	1797	4	US-09-364-230-19	Sequence 19, Appl	919	15	0.5	2178	2	US-08-463-667A-2	Sequence 2, Appl
c 847	15	0.5	1821	4	US-09-486-382B-10	Sequence 10, Appl	920	15	0.5	2178	3	US-08-923-854-24	Sequence 24, Appl
c 848	15	0.5	1834	4	US-09-105-390-4	Sequence 4, Appl	921	15	0.5	2178	5	PCT-US91-09133-25	Sequence 25, Appl
c 849	15	0.5	1835	1	US-08-361-611-3	Sequence 3, Appl	c 922	15	0.5	2262	1	US-07-689-008-3	Sequence 3, Appl
c 850	15	0.5	1835	1	US-08-565-655-3	Sequence 3, Appl	923	15	0.5	2308	4	US-09-382-256-9	Sequence 9, Appl
c 851	15	0.5	1835	2	US-08-946-967-3	Sequence 3, Appl	924	15	0.5	2308	4	US-09-395-115-9	Sequence 9, Appl
c 852	15	0.5	1841	4	US-09-053-374A-1	Sequence 1, Appl	925	15	0.5	2308	4	US-08-436-265-9	Sequence 9, Appl
c 853	15	0.5	1851	4	US-09-183-959-9	Sequence 9, Appl	926	15	0.5	2308	4	US-09-679-187-9	Sequence 9, Appl
c 854	15	0.5	1857	4	US-09-377-557-9	Sequence 9, Appl	c 927	15	0.5	2369	4	US-09-302-769-20	Sequence 20, Appl
c 855	15	0.5	1866	3	US-08-909-742-1	Sequence 1, Appl	928	15	0.5	2398	1	US-07-971-558-2	Sequence 2, Appl
c 856	15	0.5	1866	4	US-09-412-289-1	Sequence 4, Appl	929	15	0.5	2398	2	US-08-515-251A-3	Sequence 3, Appl
c 857	15	0.5	1896	4	US-09-343-011B-4	Sequence 1, Appl	930	15	0.5	2412	1	US-08-158-232-9	Sequence 9, Appl
c 858	15	0.5	1897	1	US-08-453-477-1	Sequence 1, Appl	c 931	15	0.5	2412	1	US-08-304-626-9	Sequence 9, Appl
c 859	15	0.5	1897	1	US-08-453-461-1	Sequence 1, Appl	c 932	15	0.5	2412	1	US-08-316-301A-11	Sequence 11, Appl
c 860	15	0.5	1920	4	US-09-160-496-4	Sequence 4, Appl	c 933	15	0.5	2412	1	US-08-611-928-9	Sequence 9, Appl
c 861	15	0.5	1946	2	US-08-755-584-1	Sequence 1, Appl	c 934	15	0.5	2412	3	US-09-173-891-9	Sequence 9, Appl
c 862	15	0.5	1946	3	US-09-192-611-1	Sequence 1, Appl	c 935	15	0.5	2412	3	US-09-076-137-11	Sequence 11, Appl
c 863	15	0.5	1962	1	US-08-399-696-1	Sequence 1, Appl	c 936	15	0.5	2412	5	PCT-US92-03624-11	Sequence 11, Appl
c 864	15	0.5	1984	4	US-08-937-067-5	Sequence 5, Appl	937	15	0.5	2419	4	US-09-245-041-8	Sequence 8, Appl
c 865	15	0.5	1995	3	US-08-904-452-1	Sequence 1, Appl	c 938	15	0.5	2461	1	US-08-282-141-1	Sequence 1, Appl
c 866	15	0.5	1995	4	US-09-517-639-1	Sequence 1, Appl	939	15	0.5	2481	1	US-08-324-243-35	Sequence 35, Appl
c 867	15	0.5	2018	1	US-08-188-582-6	Sequence 6, Appl	940	15	0.5	2481	1	US-08-532-390-35	Sequence 35, Appl
c 868	15	0.5	2018	1	US-08-646-715-6	Sequence 6, Appl	941	15	0.5	2481	3	US-08-717-294-35	Sequence 35, Appl
c 869	15	0.5	2025	3	US-09-160-483-2	Sequence 2, Appl	942	15	0.5	2481	5	PCT-US95-11511-35	Sequence 35, Appl
c 870	15	0.5	2031	4	US-09-149-476-155	Sequence 155, App	943	15	0.5	2505	4	US-09-097-199-85	Sequence 85, Appl
c 871	15	0.5	2033	1	US-08-148-910-14	Sequence 14, Appl	c 944	15	0.5	2505	4	US-09-268-140-1	Sequence 4, Appl
c 872	15	0.5	2033	1	US-08-448-937A-14	Sequence 14, Appl	c 945	15	0.5	2510	4	US-08-894-324A-2	Sequence 2, Appl
c 873	15	0.5	2084	1	US-08-720-899-3	Sequence 3, Appl	946	15	0.5	2511	1	US-08-363-560-1	Sequence 1, Appl
c 874	15	0.5	2084	1	US-08-459-610-3	Sequence 3, Appl	c 947	15	0.5	2517	4	US-09-268-140-7	Sequence 7, Appl
c 875	15	0.5	2084	2	US-08-343-804-3	Sequence 3, Appl	948	15	0.5	2538	3	US-08-899-437-1	Sequence 1, Appl
c 876	15	0.5	2084	2	US-08-687-399-3	Sequence 3, Appl	949	15	0.5	2538	4	US-09-126-121-1	Sequence 1, Appl
c 877	15	0.5	2084	3	US-08-600-908A-3	Sequence 3, Appl	c 950	15	0.5	2544	4	US-09-483-371-1	Sequence 1, Appl
c 878	15	0.5	2084	3	US-08-683-838A-3	Sequence 3, Appl	c 951	15	0.5	2565	4	US-09-105-058C-26	Sequence 26, Appl
c 879	15	0.5	2084	3	US-09-182-859-3	Sequence 3, Appl	952	15	0.5	2604	4	US-09-264-097-3	Sequence 3, Appl
c 880	15	0.5	2084	4	US-09-170-670-13	Sequence 13, Appl	953	15	0.5	2604	4	US-09-537-168-5	Sequence 5, Appl
c 881	15	0.5	2084	4	US-09-193-068-31	Sequence 31, Appl	954	15	0.5	2609	4	US-09-221-017B-747	Sequence 747, App
c 882	15	0.5	2084	4	US-09-183-412-58	Sequence 58, Appl	c 955	15	0.5	2629	2	US-09-012-030-1	Sequence 1, Appl
c 883	15	0.5	2084	4	US-09-290-734-13	Sequence 13, Appl	c 956	15	0.5	2629	2	US-08-590-454-1	Sequence 1, Appl
c 884	15	0.5	2084	4	US-09-672-459-3	Sequence 3, Appl	c 957	15	0.5	2639	4	US-09-374-135-1	Sequence 1, Appl
c 885	15	0.5	2084	4	US-09-636-252A-3	Sequence 3, Appl	c 958	15	0.5	2653	1	US-08-295-670-5	Sequence 5, Appl
c 886	15	0.5	2085	4	US-09-352-990-9	Sequence 9, Appl	c 959	15	0.5	2653	1	US-08-633-485-5	Sequence 5, Appl
c 887	15	0.5	2087	4	US-09-097-199-83	Sequence 83, Appl	960	15	0.5	2677	4	US-08-882-164D-36	Sequence 36, Appl
c 888	15	0.5	2094	4	US-09-106-191-8	Sequence 1, Appl	961	15	0.5	2689	2	US-08-985-090-1	Sequence 1, Appl
c 889	15	0.5	2094	4	US-09-661-753-1	Sequence 1, Appl	962	15	0.5	2689	3	US-09-165-543-1	Sequence 1, Appl
c 890	15	0.5	2108	4	US-09-032-742-6	Sequence 6, Appl	c 963	15	0.5	2693	1	US-07-872-644-50	Sequence 50, Appl
c 891	15	0.5	2114	4	US-09-568-816A-1	Sequence 1, Appl	964	15	0.5	2693	1	US-08-295-670-7	Sequence 7, Appl
c 892	15	0.5	2115	2	US-08-474-379C-60	Sequence 60, Appl	c 965	15	0.5	2693	1	US-08-297-494-50	Sequence 50, Appl
c 893	15	0.5	2115	2	US-08-474-379C-60	Sequence 60, Appl	966	15	0.5	2693	1	US-08-297-510-50	Sequence 50, Appl
c 894	15	0.5	2115	3	US-09-146-249A-60	Sequence 60, Appl	c 967	15	0.5	2693	1	US-08-633-485-7	Sequence 7, Appl
c 895	15	0.5	2115	3	US-09-146-249A-60	Sequence 60, Appl	968	15	0.5	2693	1	US-08-479-532-50	Sequence 50, Appl
c 896	15	0.5	2115	3	US-08-206-188B-60	Sequence 60, Appl	969	15	0.5	2693	1	US-08-455-526-50	Sequence 50, Appl
c 897	15	0.5	2115	3	US-08-206-188B-60	Sequence 60, Appl	970	15	0.5	2693	1	US-08-455-525-50	Sequence 50, Appl
c 898	15	0.5	2130	1	US-07-952-800-1	Sequence 1, Appl	971	15	0.5	2693	3	US-09-139-491-50	Sequence 50, Appl
c 899	15	0.5	2135	4	US-09-173-300-1	Sequence 1, Appl	972	15	0.5	2693	5	PCT-US92-03222-50	Sequence 50, Appl
c 900	15	0.5	2140	1	US-08-334-698-1	Sequence 1, Appl	c 973	15	0.5	2702	3	US-08-508-761B-3	Sequence 3, Appl
c 901	15	0.5	2140	1	US-08-228-932-1	Sequence 1, Appl	c 974	15	0.5	2742	4	US-09-232-468A-1	Sequence 1, Appl
c 902	15	0.5	2140	1	US-08-468-939-1	Sequence 1, Appl	975	15	0.5	2757	2	US-08-627-254C-26	Sequence 26, Appl
c 903	15	0.5	2140	2	US-08-406-855A-1	Sequence 1, Appl	c 976	15	0.5	2764	2	US-08-465-971B-1	Sequence 1, Appl

```
c 977      15      0.5 2787 4 US-09-624-693A-16
978      15      0.5 2817 5 PCT-US93-05944-1
979      15      0.5 2838 4 US-08-246-489-1
980      15      0.5 2853 4 US-09-221-017B-843
981      15      0.5 2870 1 US-08-468-036-28
982      15      0.5 2870 2 US-08-376-843-28
983      15      0.5 2892 1 US-08-264-534-5
984      15      0.5 2892 1 US-08-083-590A-1
985      15      0.5 2892 1 US-08-465-500-5
986      15      0.5 2892 2 US-08-346-126-5
987      15      0.5 2892 2 US-08-346-128-5
988      15      0.5 2892 3 US-08-532-384-1
989      15      0.5 2892 3 US-08-893-828-5
990      15      0.5 2897 2 US-08-927-394-1
991      15      0.5 2908 4 US-09-930-181-1
992      15      0.5 2914 4 US-09-177-650-6
993      15      0.5 2991 3 US-08-795-430-48
994      15      0.5 2991 4 US-09-353-700-48
995      15      0.5 3000 4 US-08-460-269C-3
996      15      0.5 3000 4 US-08-460-269C-5
997      15      0.5 3044 1 US-07-872-644-44
998      15      0.5 3044 1 US-08-297-494-44
999      15      0.5 3044 1 US-08-297-510-44
c1000     15      0.5 3044 1 US-08-479-532-44
```

## ALIGNMENTS

## RESULT 1

```
US-08-750-145A-23/c
; Sequence 23, Application US/08750145A
; Patent No. 6010851
; GENERAL INFORMATION:
; APPLICANT: MIHARA, Yasuhiro
; APPLICANT: UTAGAWA, Takashi
; APPLICANT: YAMADA, Hideaki
; APPLICANT: ASANO, Yasuhisa
; TITLE OF INVENTION: Method for Producing Nucleoside-5'-
; TITLE OF INVENTION: Phosphate Ester
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT, P.C.
; STREET: 1755 S. JEFF. DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/750.145A
; FILING DATE: 01-JAN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 7-149781
; FILING DATE: 05-May-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8-094680
; FILING DATE: 26-Mar-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: NORMAN F. OBLON
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 0010-0830-0 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 735 base pairs
```

```
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Serratia ficaria
; STRAIN: IAM 13540
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..732
US-08-750-145A-23

Query Match          0.6% Score 20; DB 3; Length 735;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1751 TGCAGCTGGCGCGGAAGGT 1770
      |||||||
Db 698 TGCAGCTGGCGCGGAAGGT 679

RESULT 2
US-08-975-698A-27/c
; Sequence 27, Application US/08975698A
; Patent No. 6015697
; GENERAL INFORMATION:
; APPLICANT: MIHARA, YASUHIRO
; APPLICANT: UTAGAWA, TAKASHI
; APPLICANT: YAMADA, HIDEAKI
; APPLICANT: ASANO, YASUHIKA
; TITLE OF INVENTION: METHOD FOR PRODUCING NUCLEOSIDE-5'-PHOSPHATE
; TITLE OF INVENTION: ESTER
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT, P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/975.698A
; FILING DATE: 21-NOV-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 0010-0885-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 735 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Serratia ficaria
; STRAIN: IAM 13540
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..732
```

US-08-975-698A-27

Query Match 0.6%; Score 20; DB 3; Length 735;  
Best Local Similarity 100.0%; Pred. No. 3.3;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1751 TGCAGCTGGCGCGGAAGT 1770  
|||||  
DB 698 TGCAGCTGGCGCGGAAGT 679

RESULT 3

US-09-417-090-27/c  
; Sequence 27, Application US/09417090  
; Patent No. 6207435

GENERAL INFORMATION:

APPLICANT: MIHARA, YASUHIRO  
UTAGAWA, TAKASHI  
YAMADA, HIDEAKI  
ASANO, YASUHISA

TITLE OF INVENTION: METHOD FOR PRODUCING NUCLEOSIDE-5'-PHOSPHATE  
ESTER

NUMBER OF SEQUENCES: 30

CORRESPONDENCE ADDRESS:

ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT, P.C.  
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: USA  
ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/417,090

FILING DATE: 13-Oct-1999

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/975,698

FILING DATE: 21-Nov-1997

ATTORNEY/AGENT INFORMATION:

NAME: OBLON, NORMAN F.

REGISTRATION NUMBER: 24,618

REFERENCE/DOCKET NUMBER: 0010-0885-0

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 413-3000

TELEFAX: (703) 413-2220

INFORMATION FOR SEQ ID NO: 27:

SEQUENCE CHARACTERISTICS:

LENGTH: 735 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: genomic DNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Serratia ficaria

STRAIN: IAM 13540

FEATURE:

NAME/KEY: CDS

LOCATION: 1..732

SEQUENCE DESCRIPTION: SEQ ID NO: 27:

US-09-417-090-27

Query Match 0.6%; Score 20; DB 4; Length 735;  
Best Local Similarity 100.0%; Pred. No. 3.3;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1751 TGCAGCTGGCGCGGAAGT 1770  
|||||

DB 698 TGCAGCTGGCGCGGAAGT 679

RESULT 4

US-09-727-578-27/c  
; Sequence 27, Application US/09727578  
; Patent No. 6355472

GENERAL INFORMATION:

APPLICANT: MIHARA, YASUHIRO

UTAGAWA, TAKASHI

APPLICANT: YAMADA, HIDEAKI

ASANO, YASUHISA

TITLE OF INVENTION: METHOD FOR PRODUCING NUCLEOSIDE-5'-PHOSPHATE

TITLE OF INVENTION: ESTER

NUMBER OF SEQUENCES: 30

CORRESPONDENCE ADDRESS:

ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT, P.C.  
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: USA  
ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/727,578

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/975,698

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: OBLON, NORMAN F.

REGISTRATION NUMBER: 24,618

REFERENCE/DOCKET NUMBER: 0010-0885-0

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 413-3000

TELEFAX: (703) 413-2220

INFORMATION FOR SEQ ID NO: 27:

SEQUENCE CHARACTERISTICS:

LENGTH: 735 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: genomic DNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Serratia ficaria

STRAIN: IAM 13540

FEATURE:

NAME/KEY: CDS

LOCATION: 1..732

US-09-727-578-27

Query Match 0.6%; Score 20; DB 4; Length 735;  
Best Local Similarity 100.0%; Pred. No. 3.3;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1751 TGCAGCTGGCGCGGAAGT 1770  
|||||

DB 698 TGCAGCTGGCGCGGAAGT 679

RESULT 5

US-08-219-012-74/c  
; Sequence 74, Application US/08219012  
; Patent No. 5543293

GENERAL INFORMATION:

APPLICANT: Larry Gold

APPLICANT: Diane Tasset  
TITLE OF INVENTION: Ligands of Thrombin  
NUMBER OF SEQUENCES: 92  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Beaton & Swanson, P.C.  
STREET: 4582 South Ulster Street Parkway, Suite #  
STREET: 403  
CITY: Denver  
STATE: Colorado  
COUNTRY: USA  
ZIP: 80237  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 360 Kb storage  
COMPUTER: IBM compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
FILING DATE: US/08/219,012  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA: none  
ATTORNEY/AGENT INFORMATION:  
NAME: Barry J. Swanson  
REGISTRATION NUMBER: 33,215  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 850-9900  
TELEFAX: (303) 850-9401  
INFORMATION FOR SEQ ID NO: 74:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 60 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-219-012-74  
Query Match 0.6%; Score 18; DB 1; Length 60;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 9 GACCCCAACCTTCCCTAC 26  
Db 41 GACCCCAACCTTCCCTAC 24  
RESULT 6  
US-08-687-421-262/c  
Sequence 262, Application US/08687421  
Patent No. 6177557  
GENERAL INFORMATION:  
APPLICANT: Gold, Larry  
APPLICANT: Janjic, Nebojsa  
APPLICANT: Tasset, Diane  
TITLE OF INVENTION: HIGH-AFFINITY LIGANDS OF BASIC  
TITLE OF INVENTION: FIBROBLAST GROWTH FACTOR AND  
TITLE OF INVENTION: THROMBIN  
NUMBER OF SEQUENCES: 445  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Swanson & Bratschun, L.L.C.  
STREET: 8400 E. Prentice Avenue, Suite 200  
CITY: Englewood  
STATE: Colorado  
COUNTRY: USA  
ZIP: 80111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 MB storage  
COMPUTER: IBM compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WordPerfect 6.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/687,421  
FILING DATE: 08-MAY-1996  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/195,005  
FILING DATE: 10-FEBRUARY-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE: 22-APRIL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/219,012  
FILING DATE: 28-MARCH-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/973,333  
FILING DATE: 11-NOVEMBER-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/714,131  
FILING DATE: 10-JUNE-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/536,428  
FILING DATE: 11-JUNE-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Barry J. Swanson  
REGISTRATION NUMBER: 33,215  
REFERENCE/DOCKET NUMBER: NEX07/PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 793-3333  
TELEFAX: (303) 793-3433  
INFORMATION FOR SEQ ID NO: 262:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 60 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-687-421-262  
Query Match 0.6%; Score 18; DB 4; Length 60;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 9 GACCCCAACCTTCCCTAC 26  
Db 41 GACCCCAACCTTCCCTAC 24  
RESULT 7  
US-09-221-017B-718/c  
Sequence 718, Application US/09221017B  
Patent No. 644799  
GENERAL INFORMATION:  
APPLICANT: Ross, Bruce C.  
TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF  
NUMBER OF SEQUENCES: 1120  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORRISON & FOERSTER  
STREET: 755 PAGE MILL ROAD  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304-1018  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows  
SOFTWARE: FastSEQ for Windows Version 2.0b  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/221,017B  
FILING DATE: 23-DEC-1998  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PP1182  
FILING DATE: 31-DEC-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PP1546  
FILING DATE: 30-JAN-1998  
PRIOR APPLICATION DATA:

```

; APPLICATION NUMBER: PP2911
; FILING DATE: 09-APR-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU98/01023
; FILING DATE: 10-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Monroy, Gladys H
; REGISTRATION NUMBER: 32,430
; REFERENCE/DOCKET NUMBER: 27340-20021.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-813-5600
; TELEFAX: 650-494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 718:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 692 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: UNKNOWN
; ORIGINAL SOURCE:
; ORGANISM: PORYPHYROMONAS GINGIVALIS
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1...692
; US-09-221-017B-718

```

```

Query Match 0.6%; Score 18; DB 4; Length 692;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 2118 CAATCTGGAACCTGAT 2135
Db 561 CAATCTGGAACCTGAT 544

```

```

RESULT 8
US-09-276-531-101/c
; Sequence 101, Application US/09276531
; Patent No. 6183968
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Reddy, Roopa
; APPLICANT: Guegler, Karl J.
; APPLICANT: Baughn, Mariah R.
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF GENES ENCODING
; TITLE OF INVENTION: RECEPTORS AND PROTEINS ASSOCIATED WITH CELL PROLIFERATION
; NUMBER OF SEQUENCES: 134
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/276,531
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/079,677
; FILING DATE: March 27, 1998
; CLASSIFICATION:

```

```

; ATTORNEY/AGENT INFORMATION:
; NAME: Lynn E. Murry, Ph.D.
; REGISTRATION NUMBER: 42,918
; REFERENCE/DOCKET NUMBER: PA-0008 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 101:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1689 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BLADTUT07
; CLONE: 1887573
; US-09-276-531-101
;
; Query Match 0.6%; Score 18; DB 4; Length 1689;
; Best Local Similarity 100.0%; Pred. No. 34;
; Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 2541 GGCATGGGACACAGCAC 2558
; Db 1275 GGCATGGGACACAGCAC 1258
;
; RESULT 9
; US-08-248-474-87
; Sequence 87, Application US/08248474
; Patent No. 5612471
; GENERAL INFORMATION:
; APPLICANT: MCK, BIRD, David
; APPLICANT: WILSON, Mark A.
; TITLE OF INVENTION: NEMATODE-INDUCE GENES IN TOMATO
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Khourie and Crew
; STREET: Steuart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/248,474
; FILING DATE: 25-MAY-1994
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 2307E-535
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 87:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 77 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Lycopersicon esculentum cv 'Rutgers Large
; ORGANISM: Red'
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..77
; OTHER INFORMATION: /standard_name= "DB# 239"

```

US-08-248-474-87

Query Match 0.5%; Score 17; DB 1; Length 77;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 87 AGCGCGCGCGCGGCA 103  
|||||  
Db 6 AGCGCGCGCGCGGCA 22

RESULT 10

US-08-756-849-87  
; Sequence 87, Application US/08756849  
; Patent No. 6093810

; GENERAL INFORMATION:  
; APPLICANT: Bird, David McK.  
; APPLICANT: Wilson, Mark A.  
; TITLE OF INVENTION: Nematode-Induced Genes in Tomato  
; NUMBER OF SEQUENCES: 129  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/756,849  
; FILING DATE: 26-NOV-1996

; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/248,474  
; FILING DATE: 25-MAY-1994

; ATTORNEY/AGENT INFORMATION:  
; NAME: Bastian, Kevin L.  
; REGISTRATION NUMBER: 34,774  
; REFERENCE/DOCKET NUMBER: 023070-053510US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300

; INFORMATION FOR SEQ ID NO: 87:  
; LENGTH: 77 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)

; ORIGINAL SOURCE:  
; ORGANISM: Lycopersicon esculentum cv 'Rutgers Large Red'  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: 1..77  
; OTHER INFORMATION: /standard\_name= "DB# 239"

US-08-756-849-87  
Query Match 0.5%; Score 17; DB 3; Length 77;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 87 AGCGCGCGCGCGGCA 103  
|||||  
Db 6 AGCGCGCGCGCGGCA 22

RESULT 11

US-08-906-156A-27  
; Sequence 27, Application US/08906156A

; Patent No. 6287854  
; GENERAL INFORMATION:  
; APPLICANT: SPURR, NIGEL K  
; APPLICANT: GRAY, IAN C  
; APPLICANT: STEWART, LORNA M  
; TITLE OF INVENTION: DIAGNOSIS OF SUSCEPTIBILITY TO CANCER  
; TITLE OF INVENTION: AND TREATMENT THEREOF  
; NUMBER OF SEQUENCES: 94  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHYE, P.C.  
; STREET: 1100 NORTH GLEBE ROAD  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22201

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/906,156A  
; FILING DATE: 05-AUG-1997

; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/042,655  
; FILING DATE: 02-APR-1996

; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/033,147  
; FILING DATE: 13-DEC-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/005,840

; FILING DATE: 23-OCT-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO PCT/96GB/02588  
; FILING DATE: 22-OCT-1996

; ATTORNEY/AGENT INFORMATION:  
; NAME: SADOFF, B.J.  
; REGISTRATION NUMBER: 36,663  
; REFERENCE/DOCKET NUMBER: 1090-14

; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-816-4000  
; TELEFAX: 703-816-4100  
; INFORMATION FOR SEQ ID NO: 27:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 117 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA to mRNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: HUMAN  
; US-08-906-156A-27

Query Match 0.5%; Score 17; DB 4; Length 117;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GCGCGCGCGCGGCAC 104  
|||||  
Db 59 GCGCGCGCGCGGCAC 75

RESULT 12

US-08-906-156A-28  
; Sequence 28, Application US/08906156A  
; Patent No. 6287854

; GENERAL INFORMATION:  
; APPLICANT: SPURR, NIGEL K  
; APPLICANT: GRAY, IAN C  
; APPLICANT: STEWART, LORNA M



;; TITLE OF INVENTION: DIAGNOSIS OF SUSCEPTIBILITY TO CANCER  
;; TITLE OF INVENTION: AND TREATMENT THEREOF  
;; NUMBER OF SEQUENCES: 94  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: NIXON & VANDERHYE P.C.  
;; STREET: 1100 NORTH GLEBE ROAD  
;; CITY: ARLINGTON  
;; STATE: VA  
;; COUNTRY: USA  
;; ZIP: 22201  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/906,156A  
;; FILING DATE: 05-AUG-1997  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 60/042,655  
;; FILING DATE: 02-APR-1996  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 60/033,147  
;; FILING DATE: 13-DEC-1996  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 60/005,840  
;; FILING DATE: 23-OCT-1995  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: WO PCT/96GB/02588  
;; FILING DATE: 22-OCT-1996  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: SADOFF, B.J.  
;; REGISTRATION NUMBER: 36,663  
;; REFERENCE/DOCKET NUMBER: 1090-14  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 703-816-4000  
;; TELEFAX: 703-816-4100  
;; INFORMATION FOR SEQ ID NO: 28:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 322 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: cDNA to mRNA  
;; HYPOTHETICAL: NO  
;; ANTI-SENSE: NO  
;; ORIGINAL SOURCE:  
;; ORGANISM: HUMAN  
;; US-08-906-156A-28

Query Match 0.5%; Score 17; DB 4; Length 322;  
Best Local Similarity 100.0%; Pred. No. 1.le-02;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GCGCGCGCGCGCGCAC 104  
|||||

Db 82 GCGCGCGCGCGCGCAC 98

## RESULT 13

US-09-641-638-222  
; Sequence 222, Application US/09641638  
; Patent No. 6432648

## GENERAL INFORMATION:

; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Bougueleret, Lydie  
; APPLICANT: Chumakov, Ilya

; APPLICANT: Cohen, Annick

; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING  
; TITLE OF INVENTION: GENES INVOLVED IN ARACHIDONIC ACID METABOLISM

; FILE REFERENCE: GENSET.051CPI

; CURRENT APPLICATION NUMBER: US/09/641,638

; CURRENT FILING DATE: 2000-08-16  
; PRIOR APPLICATION NUMBER: US 09/502,330  
; PRIOR FILING DATE: 2000-02-11  
; PRIOR APPLICATION NUMBER: US 60/133,200  
; PRIOR FILING DATE: 1999-05-07  
; PRIOR APPLICATION NUMBER: US 09/275,267  
; PRIOR FILING DATE: 1999-03-23  
; PRIOR APPLICATION NUMBER: US 60/119,917  
; PRIOR FILING DATE: 1999-02-12  
; NUMBER OF SEQ ID NOS: 1304  
; SOFTWARE: Patent.pm  
; SEQ ID NO 222  
; LENGTH: 362  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
; FEATURE:  
; NAME/KEY: allele  
; LOCATION: 72  
; OTHER INFORMATION: 10-87-73 : polymorphic base C or T  
; NAME/KEY: misc\_binding  
; LOCATION: 52..71  
; OTHER INFORMATION: 10-87-73.mis1, potential  
; NAME/KEY: misc\_binding  
; LOCATION: 73..92  
; OTHER INFORMATION: 10-87-73.mis2, potential complement  
; NAME/KEY: primer\_bind  
; LOCATION: 1..18  
; OTHER INFORMATION: upstream amplification primer  
; NAME/KEY: primer\_bind  
; LOCATION: 345..362  
; OTHER INFORMATION: downstream amplification primer, complement  
; NAME/KEY: misc\_binding  
; LOCATION: 60..84  
; OTHER INFORMATION: 10-87-73 potential probe  
; NAME/KEY: misc\_feature  
; LOCATION: 172..174  
; OTHER INFORMATION: n=a, g, c or t  
; US-09-641-638-222

Query Match 0.5%; Score 17; DB 4; Length 362;

Best Local Similarity 100.0%; Pred. No. 1.le-02;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 87 AGCGCGCGCGCGCA 103  
|||||

Db 24 AGCGCGCGCGCGCA 40

## RESULT 14

US-09-641-638-223

; Sequence 223, Application US/09641638

; Patent No. 6432648

## GENERAL INFORMATION:

; APPLICANT: Blumenfeld, Marta

; APPLICANT: Bougueleret, Lydie

; APPLICANT: Chumakov, Ilya

; APPLICANT: Cohen, Annick

; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING  
; TITLE OF INVENTION: GENES INVOLVED IN ARACHIDONIC ACID METABOLISM

; FILE REFERENCE: GENSET.051CPI

; CURRENT APPLICATION NUMBER: US/09/641,638

; CURRENT FILING DATE: 2000-08-16

; PRIOR APPLICATION NUMBER: US 09/502,330

; PRIOR FILING DATE: 2000-02-11

; PRIOR APPLICATION NUMBER: US 60/133,200

; PRIOR FILING DATE: 1999-05-07

; PRIOR APPLICATION NUMBER: US 09/275,267

; PRIOR FILING DATE: 1999-03-23

; PRIOR APPLICATION NUMBER: US 60/119,917

; PRIOR FILING DATE: 1999-02-12

; NUMBER OF SEQ ID NOS: 1304

; SOFTWARE: Patent.pm

; SEQ ID NO 223

```
; LENGTH: 362
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 73
; OTHER INFORMATION: 10-87-74 : polymorphic base A or T
; NAME/KEY: misc_binding
; LOCATION: 53..72
; OTHER INFORMATION: 10-87-74.misl, potential
; NAME/KEY: misc_binding
; LOCATION: 74..93
; OTHER INFORMATION: 10-87-74.mis2, potential complement
; NAME/KEY: primer_bind
; LOCATION: 1..18
; OTHER INFORMATION: upstream amplification primer
; NAME/KEY: primer_bind
; LOCATION: 345..362
; OTHER INFORMATION: downstream amplification primer, complement
; NAME/KEY: misc_binding
; LOCATION: 61..85
; OTHER INFORMATION: 10-87-74 potential probe
; NAME/KEY: misc_feature
; LOCATION: 172..174
; OTHER INFORMATION: n=a, g, c or t
US-09-641-638-223
```

```
Query Match          0.5%; Score 17; DB 4; Length 362;
Best Local Similarity 100.0%; Pred.No. 1.1e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 87 AGCGCGCGCGCGGCA 103
      |||||
Db 24 AGCGCGCGCGCGGCA 40
```

```
RESULT 15
US-09-641-638-224
; Sequence 224, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; TITLE OF INVENTION: GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
; FILE REFERENCE: GENSET.051CPI
; CURRENT APPLICATION NUMBER: US/09/641,638
; CURRENT FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 224
; LENGTH: 362
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 79
; OTHER INFORMATION: 10-87-80 : polymorphic base A or G
; NAME/KEY: misc_binding
; LOCATION: 59..78
; OTHER INFORMATION: 10-87-80.misl, potential
; NAME/KEY: misc_binding
; LOCATION: 80..99
```

```
; OTHER INFORMATION: 10-87-80.mis2, potential complement
; NAME/KEY: primer_bind
; LOCATION: 1..18
; OTHER INFORMATION: upstream amplification primer
; NAME/KEY: primer_bind
; LOCATION: 345..362
; OTHER INFORMATION: downstream amplification primer, complement
; NAME/KEY: misc_binding
; LOCATION: 67..91
; OTHER INFORMATION: 10-87-80 potential probe
; NAME/KEY: misc_feature
; LOCATION: 172..174
; OTHER INFORMATION: n=a, g, c or t
US-09-641-638-224
```

```
Query Match          0.5%; Score 17; DB 4; Length 362;
Best Local Similarity 100.0%; Pred.No. 1.1e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 87 AGCGCGCGCGCGGCA 103
      |||||
Db 24 AGCGCGCGCGCGGCA 40
```

```
RESULT 16
US-09-641-638-225
; Sequence 225, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; TITLE OF INVENTION: GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
; FILE REFERENCE: GENSET.051CPI
; CURRENT APPLICATION NUMBER: US/09/641,638
; CURRENT FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 225
; LENGTH: 362
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 138
; OTHER INFORMATION: 10-87-140 : polymorphic base C or T
; NAME/KEY: misc_binding
; LOCATION: 118..137
; OTHER INFORMATION: 10-87-140.misl, potential
; NAME/KEY: misc_binding
; LOCATION: 139..158
; OTHER INFORMATION: 10-87-140.mis2, potential complement
; NAME/KEY: primer_bind
; LOCATION: 1..18
; OTHER INFORMATION: upstream amplification primer
; NAME/KEY: primer_bind
; LOCATION: 345..362
; OTHER INFORMATION: downstream amplification primer, complement
; NAME/KEY: misc_binding
; LOCATION: 126..150
; OTHER INFORMATION: 10-87-140 potential probe
; NAME/KEY: misc_feature
; LOCATION: 172..174
```

```
; OTHER INFORMATION: n=a, g, c or t
US-09-641-638-225

Query Match      0.5%; Score 17; DB 4; Length 362;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 87 AGCGCGGGCGGGCGCA 103
    |||||
Db 24 AGCGCGGGCGGGCGCA 40

RESULT 17
US-09-060-756-653
; Sequence 653, Application US/09060756
; Patent No. 6183957
; GENERAL INFORMATION:
; APPLICANT: Cole, Stewart
; APPLICANT: Buchrieser-Brosch, Roland
; APPLICANT: Gordon, Stephen
; APPLICANT: Billault, Alain
; TITLE OF INVENTION: METHOD FOR ISOLATING A POLYNUCLEOTIDE OF INTEREST FROM
; TITLE OF INVENTION: THE GENOME OF A MYCOBACTERIUM USING A BAC-BASED DNA
; TITLE OF INVENTION: LIBRARY APPLICATION TO THE DETECTION OF MYCOBACTERIA
; FILE REFERENCE: 3495-0169
; CURRENT APPLICATION NUMBER: US/09/060,756
; CURRENT FILING DATE: 1998-04-16
; NUMBER OF SEQ ID NOS: 743
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 653
; LENGTH: 429
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (various positions within the sequence)
; OTHER INFORMATION: applicants are uncertain of bases designated as "n"
US-09-060-756-653

Query Match      0.5%; Score 17; DB 4; Length 429;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 792 TGGCGTGGCGCATCGTCA 808
    |||||
Db 202 TGGCGTGGCGCATCGTCA 218

RESULT 18
US-08-642-274D-18
; Sequence 18, Application US/08642274D
; Patent No. 6200749
; GENERAL INFORMATION:
; APPLICANT: Shiloh, Yosef
; TITLE OF INVENTION: MUTATED FORMS OF THE ATAXIA-TELANGIECTASIA GENE AND METHOD TO
; FILE REFERENCE: 229000033
; CURRENT APPLICATION NUMBER: US/08/642,274D
; CURRENT FILING DATE: 1996-05-03
; NUMBER OF SEQ ID NOS: 220
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 500
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: genomic
US-08-642-274D-18

Query Match      0.5%; Score 17; DB 4; Length 500;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 157 GCAACAACACGGAATC 173
    |||||
Db 43 GCAACAACACGGAATC 59

RESULT 19
US-08-952-014C-18
; Sequence 18, Application US/08952014C
; Patent No. 6265158
; GENERAL INFORMATION:
; APPLICANT: Shiloh, Yosef
; TITLE OF INVENTION: ATAXIA-TELANGIECTASIA GENE AND ITS
; NUMBER OF SEQUENCES: 91
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kohn & Associates
; STREET: 30500 No. 6265158thwestern Hwy., Suite 410
; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: U.S.
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/952,014C
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,995
; REFERENCE/DOCKET NUMBER: 2290.00028
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 810-539-5050
; TELEFAX: 810-539-5055
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 500 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-952-014C-18

Query Match      0.5%; Score 17; DB 4; Length 500;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 GCAACAACACGGAATC 173
    |||||
Db 43 GCAACAACACGGAATC 59

RESULT 20
US-08-998-416-826
; Sequence 826, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippssen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Reibischung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSYPPII
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
```

CITY: Research Triangle Park  
STATE: No. 6239264th Carolina  
COUNTRY: USA  
ZIP: 27709  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/998,416  
FILING DATE: 24-DEC-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: CH 0016/97  
FILING DATE: 31-DEC-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Meigs, J. Timothy  
REGISTRATION NUMBER: 38,241  
REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976  
TELEPHONE: 919-541-8587  
TELEFAX: 919-541-8689  
INFORMATION FOR SEQ ID NO: 826:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 671 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORGANISM: PAGI524UP  
US-08-998-416-826

Query Match 0.5%; Score 17; DB 4; Length 671;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2465 GCAAAATCGCGGCAGT 2481  
|||||  
Db 589 GCAAAATCGCGGCAGT 605

RESULT 21  
PCT-US95-04801-4  
Sequence 4, Application PC/TUS9504801  
GENERAL INFORMATION:  
APPLICANT: Martin, Juan F.  
APPLICANT: Coque, Juan R.  
APPLICANT: Enguita, Francisco J.  
APPLICANT: Fuente, Juan L.  
APPLICANT: Llerena, Francisco J.  
APPLICANT: Liras, Paloma  
TITLE OF INVENTION: DNA ENCODING CEPHAMYCIN BIOSYNTHESIS  
TITLE OF INVENTION: LATE GENES  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: John W. Wallen III  
STREET: P.O. Box 2000  
CITY: Rahway  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07065  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/04801  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:

NAME: Wallen III, John W.  
REGISTRATION NUMBER: 35,403  
REFERENCE/DOCKET NUMBER: 19179  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (908) 594-3905  
TELEFAX: (908) 594-4720  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 800 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
PCT-US95-04801-4  
Query Match 0.5%; Score 17; DB 5; Length 800;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 2129 ACCTGATGGTCGAACTG 2145  
|||||  
Db 307 ACCTGATGGTCGAACTG 323  
RESULT 22  
US-08-906-156A-13  
Sequence 13, Application US/08906156A  
Patent No. 6287854  
GENERAL INFORMATION:  
APPLICANT: SPURR, NIGEL K  
APPLICANT: GRAY, IAN C  
APPLICANT: STEWART, LORNA M  
TITLE OF INVENTION: DIAGNOSIS OF SUSCEPTIBILITY TO CANCER  
TITLE OF INVENTION: AND TREATMENT THEREOF  
NUMBER OF SEQUENCES: 94  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON & VANDERHUYE P.C.  
STREET: 1100 NORTH GLEBE ROAD  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22201  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/906,156A  
FILING DATE: 05-AUG-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/042,655  
FILING DATE: 02-APR-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/033,147  
FILING DATE: 13-DEC-1996  
APPLICATION NUMBER: US 60/005,840  
FILING DATE: 23-OCT-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/96GB/02588  
FILING DATE: 22-OCT-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: SADOFF, B.J.  
REGISTRATION NUMBER: 36,663  
REFERENCE/DOCKET NUMBER: 1090-14  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-816-4000  
TELEFAX: 703-816-4100  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 871 base pairs

```
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Part of gene corresponding to IMAGE 264611
US-08-906-156A-13

Query Match      0.5%; Score 17; DB 4; Length 871;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GCGCGCGCGCGCGGCAC 104
    |||
Db 112 GCGCGCGCGCGCGGCAC 128

RESULT 23
US-08-906-156A-58
; Sequence 58, Application US/08906156A
; Patent No. 6287854
; GENERAL INFORMATION:
; APPLICANT: SPURR, NIGEL K
; APPLICANT: GRAY, IAN C
; APPLICANT: STEWART, LORNA M
; TITLE OF INVENTION: DIAGNOSIS OF SUSCEPTIBILITY TO CANCER
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/906,156A
; FILING DATE: 05-AUG-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/042,655
; FILING DATE: 02-APR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/033,147
; FILING DATE: 13-DEC-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/005,840
; FILING DATE: 23-OCT-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/96GB/02588
; FILING DATE: 22-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: SADOFF, B. J.
; REGISTRATION NUMBER: 36,663
; REFERENCE/DOCKET NUMBER: 1090-14
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-816-4000
; TELEFAX: 703-816-4100
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 871 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
```

```
; ORIGINAL SOURCE:
; ORGANISM: human
; IMMEDIATE SOURCE:
; CLONE: part of IMAGE clone 264611
US-08-906-156A-58

Query Match      0.5%; Score 17; DB 4; Length 871;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GCGCGCGCGCGCGGCAC 104
    |||
Db 112 GCGCGCGCGCGCGGCAC 128

RESULT 24
US-09-267-031-5
; Sequence 5, Application US/09267031
; Patent No. 6137031
; GENERAL INFORMATION:
; APPLICANT: Zhang, Yuelin
; APPLICANT: Kinkema, Mark
; APPLICANT: Dong, Xinnian
; APPLICANT: Ronald, Pamela
; APPLICANT: Chern, MawSheng
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: DNA Binding Proteins That Interact with NPRI
; FILE REFERENCE: 023070-09250005
; CURRENT APPLICATION NUMBER: US/09/267,031
; CURRENT FILING DATE: 1999-03-11
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 925
; TYPE: DNA
; ORGANISM: Oryza sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(925)
; OTHER INFORMATION: rice bZIP gene MN38
US-09-267-031-5

Query Match      0.5%; Score 17; DB 3; Length 925;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 545 TTAAAGCTTCTTTCGAC 561
    |||
Db 527 TTAAAGCTTCTTTCGAC 543

RESULT 25
US-08-770-565-1/c
; Sequence 1, Application US/08770565
; Patent No. 5846723
; GENERAL INFORMATION:
; APPLICANT: Kim, Nam Woo
; APPLICANT: Wu, Fred
; APPLICANT: Kealey, James T.
; APPLICANT: Pruzan, Ronald
; APPLICANT: Weinrich, Scott L.
; TITLE OF INVENTION: Methods for Detecting the RNA Component of
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND and CREW LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
```

```
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/770,565
; FILING DATE: 20-DEC-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Storella, John R.
; REGISTRATION NUMBER: 32,944
; REFERENCE/DOCKET NUMBER: 015389-002300US
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 981 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc_RNA
; LOCATION: 267..715
; OTHER INFORMATION: /product= "htr"
; OTHER INFORMATION: /note= "htr transcript serves as
; OTHER INFORMATION: template in the telomerase
; OTHER INFORMATION: ribonucleoprotein"
; US-08-770-565-1
;
; Query Match 0.5%; Score 17; DB 2; Length 981;
; Best Local Similarity 100.0%; Pred. No. 1.1e+02;
; Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
QY 2188 GCGGCCGCGGACCGCAC 2204
DB 231 GCGGCCGCGGACCGCAC 215

; RESULT 26
; US-08-710-249-5/c
; Sequence 5, Application US/08710249
; Patent No. 5858777
; GENERAL INFORMATION:
; APPLICANT: Villeponteau, Bryant
; APPLICANT: Feng, Junli
; APPLICANT: Andrews, William H.
; APPLICANT: Adams, Robert R.
; TITLE OF INVENTION: Methods and Reagents for Regulating
; TITLE OF INVENTION: Telomere Length and Telomerase Activity
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/710,249
; FILING DATE: 13-SEP-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/583,808
; FILING DATE: 05-JAN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/003,492
; FILING DATE: 08-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Storella, John R.
; REGISTRATION NUMBER: 32,944
; REFERENCE/DOCKET NUMBER: 015389-001220US
; TELEPHONE: (415) 576-0200
;
; Query Match 0.5%; Score 17; DB 2; Length 981;
; Best Local Similarity 100.0%; Pred. No. 1.1e+02;
; Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
QY 2188 GCGGCCGCGGACCGCAC 2204
DB 231 GCGGCCGCGGACCGCAC 215

; RESULT 27
; US-08-833-377-1/c
; Sequence 1, Application US/08833377
; Patent No. 5968506
; GENERAL INFORMATION:
; APPLICANT: Weinrich, Scott L.
; APPLICANT: Atkinson III, Edward M.
; APPLICANT: Lichtsteiner, Serge P.
; APPLICANT: Vasserot, Alain P.
; APPLICANT: Pruzan, Ronald A.
; APPLICANT: Kealey, James T.
; TITLE OF INVENTION: Purified Telomerase
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/833,377
; FILING DATE: 04-APR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/510,736
; FILING DATE: 04-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Storella, John R.
; REGISTRATION NUMBER: 32,944
; REFERENCE/DOCKET NUMBER: 015389-001110US
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 981 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
```

NAME/KEY: -  
LOCATION: 1..981  
OTHER INFORMATION: /note= "PstI fragment of the 2.4 kb  
SauIIIA1-HindIII fragment of clone 28-1"  
US-08-833-377-1

Query Match 0.5%; Score 17; DB 2; Length 981;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2188 GCGGCGCGCGACCGCAC 2204  
|||||  
Db 231 GCGGCGCGCGACCGCAC 215

RESULT 28  
US-08-714-482-2/c  
Sequence 2, Application US/08714482  
Patent No. 5972605  
GENERAL INFORMATION:  
APPLICANT: Villeeponteau, Bryant  
APPLICANT: Harley, Calvin  
TITLE OF INVENTION: Assays for Regulators of Mammalian  
Telomerase Expression  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/714,482  
FILING DATE: 16-SEP-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/521,634  
FILING DATE: 31-AUG-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/482,115  
FILING DATE: 07-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/472,802  
FILING DATE: 07-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/330,123  
FILING DATE: 27-OCT-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/272,102  
FILING DATE: 07-JUL-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Storella, John R.  
REGISTRATION NUMBER: 32,944  
REFERENCE/DOCKET NUMBER: 015389-008600S  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 981 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: -  
LOCATION: 1..981

OTHER INFORMATION: /note= "PstI fragment containing htr  
sequence"  
US-08-714-482-2

Query Match 0.5%; Score 17; DB 2; Length 981;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2188 GCGGCGCGCGACCGCAC 2204  
|||||  
Db 231 GCGGCGCGCGACCGCAC 215

RESULT 29  
US-08-838-545-22/c  
Sequence 22, Application US/08838545  
Patent No. 6046307  
GENERAL INFORMATION:  
APPLICANT: Shay, Jerry W.  
APPLICANT: Wright, Woodring E.  
APPLICANT: Piatyszek, Mieczyslaw A.  
APPLICANT: Corey, David R.  
APPLICANT: No. 6046307ton, James C.  
TITLE OF INVENTION: Modulation of Mammalian Telomerase by  
Peptide Nucleic Acids  
NUMBER OF SEQUENCES: 60  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/838,545  
FILING DATE: 09-APR-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/630,019  
FILING DATE: 09-APR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Storella, John R.  
REGISTRATION NUMBER: 32,944  
REFERENCE/DOCKET NUMBER: 015389-0016100S  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 981 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: misc\_RNA  
LOCATION: 266..716  
OTHER INFORMATION: /product= "htr"  
OTHER INFORMATION: /note= "RNA component of human telomerase (htr)"  
FEATURE:  
NAME/KEY: -  
LOCATION: 1..981  
OTHER INFORMATION: /note= "PstI fragment of  
SauIIIA1-HindIII fragment of clone 28-1"  
US-08-838-545-22

Query Match 0.5%; Score 17; DB 3; Length 981;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2188 GCGGCCCGCCGACCGCAC 2204  
|||||

Db 231 GCGGCCCGCCGACCGCAC 215

## RESULT 30

US-09-349-532-22/c  
; Sequence 22, Application US/09349532  
; Patent No. 6294650  
; GENERAL INFORMATION:  
; APPLICANT: Shay, Jerry W.  
; APPLICANT: Wright, Woodring E.  
; APPLICANT: Piatyszek, Mieczyslaw A.  
; APPLICANT: Corey, David R.  
; APPLICANT: No. 6294650ton, James C.  
; TITLE OF INVENTION: Modulation of Mammalian Telomerase by  
; TITLE OF INVENTION: Peptide Nucleic Acids  
; NUMBER OF SEQUENCES: 60  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/349,532  
; FILING DATE:

CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/838,545  
; FILING DATE: 09-APR-1997  
; APPLICATION NUMBER: US 08/630,019  
; FILING DATE: 09-APR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Storella, John R.  
; REGISTRATION NUMBER: 32,944  
; REFERENCE/DOCKET NUMBER: 015389-001610US  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 22:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 981 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: misc\_RNA  
; LOCATION: 266..716  
; OTHER INFORMATION: /product= "htr"  
; OTHER INFORMATION: /note= "RNA component of human telomerase (htr)"  
; FEATURE:  
; NAME/KEY: -  
; LOCATION: 1..981  
; OTHER INFORMATION: /note= "PstI fragment of  
; OTHER INFORMATION: SauIII-HindIII fragment of clone 28-1"

Query Match 0.5%; Score 17; DB 4; Length 981;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2188 GCGGCCCGCCGACCGCAC 2204  
|||||

Db 231 GCGGCCCGCCGACCGCAC 215

## RESULT 31

US-09-220-157A-5/C  
; Sequence 5, Application US/09220157A  
; Patent No. 6300110  
; GENERAL INFORMATION:  
; APPLICANT: Villeponteau, Bryant  
; APPLICANT: Feng, Junli  
; APPLICANT: Andrews, William H.  
; APPLICANT: Adams, Robert R.  
; TITLE OF INVENTION: Methods and Reagents for Regulating  
; TITLE OF INVENTION: Telomere Length and Telomerase Activity  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/220,157A  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/710,249  
; FILING DATE: 13-SEP-1996  
; APPLICATION NUMBER: US 08/583,808  
; FILING DATE: 05-JAN-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/003,492  
; FILING DATE: 08-SEP-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Storella, John R.  
; REGISTRATION NUMBER: 32,944  
; REFERENCE/DOCKET NUMBER: 015389-001220US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 981 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: misc\_RNA  
; LOCATION: 267..715  
; OTHER INFORMATION: /product= "htr"  
; OTHER INFORMATION: /note= "htr transcript serves as  
; OTHER INFORMATION: template in the telomerase  
; OTHER INFORMATION: ribonucleoprotein"  
; US-09-220-157A-5

Query Match 0.5%; Score 17; DB 4; Length 981;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2188 GCGGCCCGCCGACCGCAC 2204  
|||||

Db 231 GCGGCCCGCCGACCGCAC 215

## RESULT 32

US-07-601-094-30/c



```

; Sequence 30, Application US/07601094
; Patent No. 5215892
; GENERAL INFORMATION:
; APPLICANT: Kishimoto, Tadimitsu
; APPLICANT: Hirano, Toshio
; APPLICANT: Akira, Shizuo
; APPLICANT: Isshiki, Hiroshi
; APPLICANT: Tanabe, Osamu
; APPLICANT: Kinoshita, Shigemi
; APPLICANT: Shimamoto, Takuya
; TITLE OF INVENTION: C/EBP2 Gene and Recombinant
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sughrue, Mion, Zinn, Macpeak &
; ADDRESSEE: Seas
; STREET: 2100 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States
; ZIP: 20037-3202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/601,094
; FILING DATE: 19901022
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1035 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1035
; TELEPHONE: (202) 293-7860
; TELEFAX: (202) 293-7860
; TELEX: 6491103
; US-07-601-094-30

Query Match 0.5%; Score 17; DB 1; Length 1035;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 87 AGCGGGCGCGCGGCA 103
Db 57 AGCGGGCGCGCGGCA 41

RESULT 33
US-08-012-735-30/c
; Sequence 30, Application US/08012735
; Patent No. 5360894
; GENERAL INFORMATION:
; APPLICANT: Kishimoto, Tadimitsu
; APPLICANT: Hirano, Toshio
; APPLICANT: Akira, Shizuo
; APPLICANT: Isshiki, Hiroshi
; APPLICANT: Tanabe, Osamu
; APPLICANT: Kinoshita, Shigemi
; APPLICANT: Shimamoto, Takuya
; TITLE OF INVENTION: C/EBP2 Gene and Recombinant
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sughrue, Mion, Zinn, Macpeak &

```

```

; ADDRESSEE: Seas
; STREET: 2100 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States
; ZIP: 20037-3202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/012,735
; FILING DATE: 19930203
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/601,094
; FILING DATE: 22 OCT 1990
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 293-7060
; TELEFAX: (202) 293-7860
; TELEX: 6491103
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1035 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1035
; OTHER INFORMATION:
; US-08-012-735-30

Query Match 0.5%; Score 17; DB 1; Length 1035;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 87 AGCGGGCGCGCGGCA 103
Db 57 AGCGGGCGCGCGGCA 41

RESULT 34
US-08-791-115B-8
; Sequence 8, Application US/08791115B
; Patent No. 6262242
; GENERAL INFORMATION:
; APPLICANT: Steck, Peter
; APPLICANT: Pershouse, Mark A.
; APPLICANT: Jasser, Samar
; APPLICANT: Yung, W.K. Alfred
; APPLICANT: Tavtligian, Sean V.
; TITLE OF INVENTION: A TUMOR SUPPRESSOR DESIGNATED TS10Q23.3
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rothwell, Figg, Ernst & Kurz, P.C.
; STREET: 555 Thirteenth Street, N.W., Suite 701-E
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 22204
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/791,115B
; FILING DATE: 30-JAN-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:

```

NAME: Ihnen, Jeffrey L.  
REGISTRATION NUMBER: 38,957  
REFERENCE/DOCKET NUMBER: 2318-134.A  
TELEPHONE: 202-683-6040  
TELEFAX: 202-683-7031  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1257 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-791-115B-8

Query Match 0.5%; Score 17; DB 4; Length 1257;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GCGCGCGCGCGGCAC 104  
|||||  
DB 490 GCGCGCGCGCGGCAC 506

RESULT 35  
US-08-224-482-5/c  
Sequence 5, Application US/08224482  
Patent No. 5837692  
GENERAL INFORMATION:  
APPLICANT: Marcola, Dan  
APPLICANT: Adamson, Eileen D.  
TITLE OF INVENTION: Inhibition of the Mitogenic Activity of  
TITLE OF INVENTION: PDGF by Mammalian EGR  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Campbell and Flores  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/224,482  
FILING DATE: 07-APR-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-ME 9913  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1440 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 57..1422  
US-08-224-482-5

Query Match 0.5%; Score 17; DB 2; Length 1440;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 TAGCGCGCGCGCGGC 102  
|||||

DB 982 TAGCGCGCGCGCGGC 966

RESULT 36  
US-08-906-156A-11  
Sequence 11, Application US/08906156A  
Patent No. 6287854  
GENERAL INFORMATION:  
APPLICANT: SPURR, NIGEL K  
APPLICANT: GRAY, IAN C  
APPLICANT: STEWART, LORNA M  
TITLE OF INVENTION: DIAGNOSIS OF SUSCEPTIBILITY TO CANCER  
TITLE OF INVENTION: AND TREATMENT THEREOF  
NUMBER OF SEQUENCES: 94  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON & VANDERHYE P.C.  
STREET: 1100 NORTH GLEBE ROAD  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22201  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/906,156A  
FILING DATE: 05-AUG-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/042,655  
FILING DATE: 02-APR-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/033,147  
FILING DATE: 13-DEC-1996  
APPLICATION NUMBER: US 60/005,840  
FILING DATE: 23-OCT-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/96GB/02588  
FILING DATE: 22-OCT-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: SADOFF, B.J.  
REGISTRATION NUMBER: 36,663  
REFERENCE/DOCKET NUMBER: 1090-14  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-816-4000  
TELEFAX: 703-816-4100  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1773 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Partial cDNA of gene corresponding to IMAGE 264611  
US-08-906-156A-11

Query Match 0.5%; Score 17; DB 4; Length 1773;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GCGCGCGCGCGGCAC 104  
|||||  
DB 112 GCGCGCGCGCGGCAC 128

RESULT 37  
US-09-310-363C-7/c  
Sequence 7, Application US/09310363C

Patent No. 6388169  
GENERAL INFORMATION:  
APPLICANT: Mahajan, Pramod B.  
APPLICANT: Shi, Jinrui  
APPLICANT: McElver, John  
APPLICANT: Bowen, Benjamin  
APPLICANT: Baszczyński, Christopher  
TITLE OF INVENTION: RecA cDNAs and Uses Thereof  
FILE REFERENCE: 0855  
CURRENT APPLICATION NUMBER: US/09/310,363C  
CURRENT FILING DATE: 1999-05-12  
PRIOR APPLICATION NUMBER: 60/099,765  
PRIOR FILING DATE: 1998-09-10  
PRIOR APPLICATION NUMBER: 60/096,492  
PRIOR FILING DATE: 1998-08-14  
PRIOR APPLICATION NUMBER: 60/088,529  
PRIOR FILING DATE: 1998-06-08  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 7  
LENGTH: 1773  
TYPE: DNA  
ORGANISM: Zea mays  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (97)...(1371)  
US-09-310-363C-7

Query Match 0.5%; Score 17; DB 4; Length 1773;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 87 AGCGCGCGCGCGGCA 103  
|||||  
Db 154 AGCGCGCGCGCGGCA 138

RESULT 38  
PCT-US95-02455-1  
Sequence 1, Application PC/TUS9502455  
GENERAL INFORMATION:  
APPLICANT: Jacobs et al  
TITLE OF INVENTION: GENE FOR MYCOBACTERIAL  
DIAMINOPIMELIC ACID  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Amster, Rothstein & Ebenstein  
STREET: 90 Park Avenue  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10016  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch 1.44 Mb  
MEDIUM TYPE: storage diskette  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Word Processor (ASCII)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/02455  
FILING DATE: Not Yet Assigned  
CLASSIFICATION:  
PRIOR APPLICATION DATA: none  
APPLICATION NUMBER: none  
FILING DATE: none  
ATTORNEY/AGENT INFORMATION:  
NAME: George, Kenneth P.  
REGISTRATION NUMBER: 30,259  
REFERENCE/DOCKET NUMBER: 96700/358  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 697-5995  
TELEFAX: (212) 286-0854 or 286-0082  
TELEX: TWX 710-581-4766

INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1801  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE:  
DESCRIPTION: dapB gene  
HYPOTHETICAL: no  
ANTI-SENSE:  
FRAGMENT TYPE:  
ORIGINAL SOURCE: BCG  
ORGANISM: BCG  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE:  
CELL LINE:  
ORGANELLE:  
IMMEDIATE SOURCE:  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT:  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
PUBLICATION INFORMATION: none  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:  
PCT-US95-02455-1

Query Match 0.5%; Score 17; DB 5; Length 1801;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 88 GCGCGCGCGCGGCAC 104  
|||||  
Db 663 GCGCGCGCGCGGCAC 679

RESULT 39  
US-09-593-711A-3/c  
Sequence 3, Application US/09593711A  
Patent No. 6271030  
GENERAL INFORMATION:  
APPLICANT: Brett P. Monia  
APPLICANT: Madeline M. Butler  
APPLICANT: Jacqueline Wyatt  
TITLE OF INVENTION: ANTISENSE MODULATION OF C/EBP BETA EXPRESSION  
FILE REFERENCE: RTS-0118  
CURRENT APPLICATION NUMBER: US/09/593,711A  
CURRENT FILING DATE: 2000-06-14  
NUMBER OF SEQ ID NOS: 244  
SEQ ID NO 3  
LENGTH: 1910  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION:  
NAME/KEY: unsure  
LOCATION: 1415  
OTHER INFORMATION: unknown

```
; NAME/KEY: unsure
; LOCATION: 1421
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 1422
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 1423
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 1424
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 1458
; OTHER INFORMATION: unknown
; NAME/KEY: CDS
; LOCATION: (299)...(1336)
US-09-593-711A-3

Query Match          0.5%; Score 17; DB 4; Length 1910;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  87 AGCGCGCGCGCGGCA 103
    |||||
Db   355 AGCGCGCGCGCGGCA 339
    |||||

RESULT 40
US-07-601-094-1/c
; Sequence 1, Application US/07601094
; Patent No. 5215892
; GENERAL INFORMATION:
; APPLICANT: Kishimoto, Tadimitsu
; APPLICANT: Hirano, Toshio
; APPLICANT: Akira, Shizuo
; APPLICANT: Ishiki, Hiroshi
; APPLICANT: Tanabe, Osamu
; APPLICANT: Kinoshita, Shigemi
; APPLICANT: Shimamoto, Takuya
; TITLE OF INVENTION: C/EBP2 Gene and Recombinant
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seas
; STREET: 2100 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States
; ZIP: 20037-3202
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/601,094
; FILING DATE: 19901022
; CLASSIFICATION: 435
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 293-7060
; TELEFAX: (202) 293-7860
; TELEX: 6491103
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1914 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
```

```
; LOCATION: 281..1316
; OTHER INFORMATION:
US-07-601-094-1

Query Match          0.5%; Score 17; DB 1; Length 1914;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  87 AGCGCGCGCGCGGCA 103
    |||||
Db   337 AGCGCGCGCGCGGCA 321
    |||||

Search completed: January 27, 2003, 23:07:26
Job time : 163 secs
```

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 27, 2003, 23:07:29 ; Search time 104 Seconds  
(without alignments)  
13840.812 Million cell updates/sec

Title: US-09-830-433a-7  
Perfect score: 3204  
Sequence: 1 atgcgaacgacccaacctt.....gcgtaggctaccggttctga 3204

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 396772 seqs, 224632407 residues

Word size : 0

Total number of hits satisfying chosen parameters: 793544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Published\_Applications\_NA:\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*

2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*

3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*

4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*

5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*

6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*

7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*

8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*

9: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*

10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*

11: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*

12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*

13: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*

14: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	177	5.5	286	9	US-09-928-457-95
2	19	0.6	284	10	US-09-867-701-2325
3	19	0.6	1267	10	US-09-925-299-125
4	19	0.6	1551	10	US-09-880-107-2355
5	19	0.6	1987	10	US-09-925-301-306
6	18	0.6	213	10	US-09-974-300-4730
7	18	0.6	276	10	US-09-878-574-15701
8	18	0.6	495	10	US-09-728-445-822
9	18	0.6	1044	10	US-09-974-300-2315
10	18	0.6	1169	9	US-10-098-841-169
11	18	0.6	2321	9	US-10-071-766-34
12	18	0.6	4446	10	US-09-815-242-7983
13	18	0.6	29729	10	US-09-070-927A-238
14	17	0.5	102	10	US-09-789-836-22
15	17	0.5	105	10	US-09-789-836-21
16	17	0.5	106	10	US-09-789-836-33
17	17	0.5	122	10	US-09-783-590-5143
18	17	0.5	126	10	US-09-974-300-4131
19	17	0.5	224	10	US-09-878-574-8654



239	16	0.5	511	10	US-09-939-980-182	Sequence 182, App	c 312	16	0.5	1045	12	US-10-062-254-161	Sequence 161, App
240	16	0.5	511	10	US-09-738-973-359	Sequence 359, App	313	16	0.5	1071	10	US-09-815-242-7426	Sequence 7426, Ap
c 241	16	0.5	514	9	US-09-933-797-761	Sequence 761, App	c 314	16	0.5	1080	10	US-09-919-580-909	Sequence 909, App
242	16	0.5	521	10	US-09-864-761-8078	Sequence 8078, Ap	315	16	0.5	1089	10	US-09-815-242-7255	Sequence 7255, Ap
c 243	16	0.5	523	10	US-09-864-761-8631	Sequence 8631, Ap	c 316	16	0.5	1091	10	US-09-822-849A-388	Sequence 388, App
c 244	16	0.5	545	10	US-09-919-580-587	Sequence 587, App	c 317	16	0.5	1113	12	US-10-062-254-205	Sequence 205, App
245	16	0.5	546	10	US-09-998-598-1151	Sequence 1151, Ap	c 318	16	0.5	1128	10	US-09-764-878-426	Sequence 426, App
246	16	0.5	547	10	US-09-867-701-1135	Sequence 1135, Ap	c 319	16	0.5	1161	9	US-09-910-186A-25	Sequence 25, Appl
c 247	16	0.5	550	10	US-09-919-580-614	Sequence 614, App	c 320	16	0.5	1161	10	US-09-833-381-2046	Sequence 2046, Ap
c 248	16	0.5	564	9	US-10-040-739-1008	Sequence 1008, Ap	c 321	16	0.5	1176	10	US-09-815-242-7814	Sequence 7814, Ap
249	16	0.5	567	10	US-09-867-701-2010	Sequence 2010, Ap	c 322	16	0.5	1179	10	US-09-912-020-188	Sequence 188, App
c 250	16	0.5	567	10	US-09-864-761-8352	Sequence 8352, Ap	c 323	16	0.5	1187	9	US-09-938-842A-5303	Sequence 5303, Ap
c 251	16	0.5	567	10	US-09-919-580-605	Sequence 605, App	c 324	16	0.5	1200	10	US-09-942-572-1	Sequence 1, Appli
252	16	0.5	577	10	US-09-864-761-12956	Sequence 12956, A	c 325	16	0.5	1210	9	US-09-764-868-93	Sequence 93, Appl
253	16	0.5	577	10	US-09-867-701-2139	Sequence 2139, Ap	c 326	16	0.5	1212	10	US-09-880-503-15	Sequence 15, Appl
c 254	16	0.5	578	10	US-09-864-761-16422	Sequence 16422, A	c 327	16	0.5	1224	10	US-09-804-551B-23	Sequence 23, Appl
c 255	16	0.5	579	10	US-09-864-761-13007	Sequence 13007, A	c 328	16	0.5	1230	10	US-09-770-44-35	Sequence 35, Appl
256	16	0.5	588	9	US-10-164-871-1	Sequence 1, Appli	c 329	16	0.5	1235	10	US-09-974-300-7279	Sequence 7279, Ap
c 257	16	0.5	593	10	US-09-919-580-504	Sequence 504, App	c 330	16	0.5	1236	10	US-09-880-503-12	Sequence 12, Appl
c 258	16	0.5	599	10	US-09-864-761-7244	Sequence 7244, Ap	c 331	16	0.5	1254	10	US-09-965-553-25	Sequence 25, Appl
c 259	16	0.5	599	10	US-09-864-761-8950	Sequence 8950, Ap	c 332	16	0.5	1260	10	US-09-175-254-2	Sequence 2, Appli
c 260	16	0.5	602	10	US-09-770-149-955	Sequence 955, App	c 333	16	0.5	1260	10	US-09-870-379-1	Sequence 1, Appli
261	16	0.5	602	10	US-09-772-134B-34	Sequence 34, Appl	c 334	16	0.5	1263	9	US-09-738-626-281	Sequence 281, App
262	16	0.5	602	10	US-09-772-134B-75	Sequence 75, Appl	c 335	16	0.5	1277	12	US-10-062-254-279	Sequence 279, App
c 263	16	0.5	603	9	US-09-764-904-72	Sequence 72, Appl	c 336	16	0.5	1283	10	US-09-925-302-307	Sequence 307, App
c 264	16	0.5	603	10	US-09-764-860-600	Sequence 600, App	c 337	16	0.5	1287	10	US-09-741-669-201	Sequence 201, App
c 265	16	0.5	607	10	US-09-864-864-242	Sequence 242, App	c 338	16	0.5	1293	10	US-09-815-242-8128	Sequence 8128, Ap
266	16	0.5	643	10	US-09-815-343-408	Sequence 408, App	c 339	16	0.5	1293	10	US-09-815-242-8727	Sequence 8727, Ap
267	16	0.5	644	10	US-09-772-134B-35	Sequence 35, Appl	c 340	16	0.5	1293	10	US-09-815-242-8952	Sequence 8952, Ap
c 268	16	0.5	646	10	US-09-867-550-1769	Sequence 1769, Ap	c 341	16	0.5	1294	9	US-09-850-964-4	Sequence 4, Appli
c 269	16	0.5	649	10	US-09-919-580-29	Sequence 29, Appl	c 342	16	0.5	1296	9	US-10-076-421-1	Sequence 1, Appli
c 270	16	0.5	649	10	US-09-772-134B-16	Sequence 16, Appl	c 343	16	0.5	1305	10	US-09-815-242-4486	Sequence 4486, Ap
c 271	16	0.5	654	10	US-09-813-358-10	Sequence 10, Appl	c 344	16	0.5	1321	9	US-09-934-249-1	Sequence 1, Appli
272	16	0.5	661	10	US-09-772-134B-86	Sequence 86, Appl	c 345	16	0.5	1333	9	US-09-841-730-24	Sequence 24, Appl
c 273	16	0.5	704	10	US-09-950-688-1	Sequence 1, Appli	c 346	16	0.5	1393	10	US-09-454-540-1	Sequence 1, Appli
c 274	16	0.5	707	9	US-10-098-841-19	Sequence 19, Appl	c 347	16	0.5	1393	10	US-09-859-894A-1	Sequence 1, Appli
c 275	16	0.5	728	10	US-09-764-898-78	Sequence 78, Appl	c 348	16	0.5	1403	9	US-10-098-841-280	Sequence 280, App
c 276	16	0.5	735	10	US-09-728-446-721	Sequence 721, App	c 349	16	0.5	1412	10	US-09-974-300-171	Sequence 171, App
c 277	16	0.5	739	10	US-09-752-639-2	Sequence 2, Appli	c 350	16	0.5	1422	9	US-09-738-626-3162	Sequence 3162, Ap
c 278	16	0.5	739	10	US-09-984-198-2	Sequence 2, Appli	c 351	16	0.5	1434	9	US-09-738-626-383	Sequence 383, App
c 279	16	0.5	740	10	US-09-974-300-4231	Sequence 4231, Ap	c 352	16	0.5	1434	9	US-10-047-412A-9	Sequence 9, Appli
c 280	16	0.5	744	9	US-09-938-842A-1922	Sequence 1922, Ap	c 353	16	0.5	1434	9	US-10-047-412A-28	Sequence 28, Appl
c 281	16	0.5	747	10	US-09-974-300-2802	Sequence 2802, Ap	c 354	16	0.5	1443	9	US-09-974-300-2435	Sequence 2435, Ap
282	16	0.5	768	9	US-09-971-536-14	Sequence 14, Appl	c 355	16	0.5	1457	9	US-09-954-531-982	Sequence 982, App
283	16	0.5	774	9	US-09-738-626-1292	Sequence 1292, Ap	c 356	16	0.5	1475	10	US-09-735-705-122	Sequence 122, App
c 284	16	0.5	776	10	US-09-919-580-889	Sequence 889, App	c 357	16	0.5	1475	10	US-09-850-718A-122	Sequence 122, App
c 285	16	0.5	791	10	US-09-919-580-912	Sequence 912, App	c 358	16	0.5	1476	9	US-09-897-778-122	Sequence 122, App
c 286	16	0.5	794	10	US-09-772-134B-18	Sequence 18, Appl	c 359	16	0.5	1511	9	US-10-078-650-11	Sequence 11, Appl
c 287	16	0.5	825	12	US-10-062-254-163	Sequence 163, App	c 360	16	0.5	1519	10	US-09-392-822-2	Sequence 2, Appli
c 288	16	0.5	831	10	US-09-880-503-14	Sequence 14, Appl	c 361	16	0.5	1523	9	US-09-938-842A-4708	Sequence 4708, Ap
c 289	16	0.5	847	9	US-09-890-813-11	Sequence 11, Appl	c 362	16	0.5	1529	10	US-09-820-893-20	Sequence 20, Appl
c 290	16	0.5	860	10	US-09-867-550-1249	Sequence 1249, Ap	c 363	16	0.5	1604	9	US-09-938-842A-3688	Sequence 3688, Ap
c 291	16	0.5	861	10	US-09-772-134B-20	Sequence 20, Appl	c 364	16	0.5	1635	9	US-10-098-841-207	Sequence 207, App
c 292	16	0.5	885	10	US-09-974-300-5111	Sequence 5111, Ap	c 365	16	0.5	1636	10	US-09-954-456-1185	Sequence 1185, Ap
c 293	16	0.5	887	7	US-08-841-636A-36	Sequence 36, Appl	c 366	16	0.5	1636	10	US-09-954-456-1850	Sequence 1850, Ap
c 294	16	0.5	903	10	US-09-919-580-927	Sequence 927, App	c 367	16	0.5	1674	9	US-09-925-301-353	Sequence 353, App
c 295	16	0.5	906	10	US-09-815-242-7590	Sequence 7590, Ap	c 368	16	0.5	1674	9	US-09-935-726-1	Sequence 1, Appli
c 296	16	0.5	920	10	US-09-765-231A-3	Sequence 3, Appli	c 369	16	0.5	1674	12	US-10-127-551-1	Sequence 1, Appli
c 297	16	0.5	937	12	US-10-062-254-165	Sequence 165, App	c 370	16	0.5	1686	9	US-10-098-841-206	Sequence 206, App
c 298	16	0.5	939	10	US-09-974-300-572	Sequence 572, App	c 371	16	0.5	1693	10	US-09-867-550-1669	Sequence 1669, Ap
c 299	16	0.5	954	9	US-09-738-626-32	Sequence 32, Appl	c 372	16	0.5	1695	10	US-09-764-878-428	Sequence 428, App
c 300	16	0.5	954	9	US-09-738-626-183	Sequence 183, App	c 373	16	0.5	1707	10	US-09-804-357-13	Sequence 13, Appl
c 301	16	0.5	957	10	US-09-822-830A-259	Sequence 259, App	c 374	16	0.5	1707	10	US-09-804-006-13	Sequence 13, Appl
c 302	16	0.5	972	10	US-09-880-503-16	Sequence 16, Appl	c 375	16	0.5	1714	10	US-09-880-192-17	Sequence 17, Appl
c 303	16	0.5	972	10	US-09-749-728B-10	Sequence 10, Appl	c 376	16	0.5	1720	10	US-09-746-359A-52	Sequence 52, Appl
c 304	16	0.5	1002	12	US-10-062-254-203	Sequence 203, App	c 377	16	0.5	1722	12	US-10-062-254-271	Sequence 271, App
305	16	0.5	1006	9	US-09-946-807-104	Sequence 104, App	c 378	16	0.5	1732	10	US-09-923-556-1	Sequence 1, Appli
306	16	0.5	1006	9	US-09-795-668-104	Sequence 104, App	c 379	16	0.5	1732	10	US-09-923-556-5	Sequence 5, Appli
307	16	0.5	1006	9	US-09-795-668-104	Sequence 104, App	c 380	16	0.5	1740	9	US-09-922-364A-15	Sequence 15, Appl
308	16	0.5	1017	9	US-09-738-626-485	Sequence 485, App	c 381	16	0.5	1740	9	US-09-254-590-15	Sequence 15, Appl
c 309	16	0.5	1020	9	US-09-712-363-53	Sequence 53, Appl	c 382	16	0.5	1740	9	US-10-115-695-15	Sequence 15, Appl
310	16	0.5	1032	10	US-09-771-161A-79	Sequence 79, Appl	c 383	16	0.5	1749	9	US-09-938-842A-1851	Sequence 1851, Ap
311	16	0.5	1035	10	US-09-866-582-5	Sequence 5, Appli	c 384	16	0.5				

385	16	0.5	1750	9	US-10-028-072-187	Sequence 187, App	458	16	0.5	2470	10	US-09-964-824A-555	Sequence 555, App
386	16	0.5	1750	9	US-10-174-590-397	Sequence 397, App	459	16	0.5	2470	10	US-09-880-107-3310	Sequence 3310, App
387	16	0.5	1750	9	US-10-176-758-397	Sequence 397, App	c 460	16	0.5	2475	9	US-10-011-588-32	Sequence 32, Appl
388	16	0.5	1750	9	US-10-175-737-397	Sequence 397, App	461	16	0.5	2475	9	US-10-028-072-195	Sequence 195, App
389	16	0.5	1750	12	US-10-052-586-397	Sequence 397, App	462	16	0.5	2475	10	US-09-815-028-1	Sequence 1, Appli
c 390	16	0.5	1779	10	US-09-833-381-1466	Sequence 1466, App	c 463	16	0.5	2543	12	US-10-044-090-211	Sequence 211, App
c 391	16	0.5	1779	10	US-09-833-381-1471	Sequence 1471, App	c 464	16	0.5	2565	9	US-09-754-853A-5	Sequence 5, Appli
c 392	16	0.5	1785	9	US-09-981-876-112	Sequence 112, App	c 465	16	0.5	2565	9	US-09-754-853A-36	Sequence 36, Appl
c 393	16	0.5	1791	10	US-09-880-107-3385	Sequence 3385, App	c 466	16	0.5	2565	9	US-09-754-853A-37	Sequence 37, Appl
c 394	16	0.5	1792	9	US-09-954-531-1349	Sequence 1349, App	c 467	16	0.5	2565	9	US-09-754-853A-38	Sequence 38, Appl
c 395	16	0.5	1799	12	US-10-062-254-247	Sequence 247, App	c 468	16	0.5	2565	9	US-09-754-853A-39	Sequence 39, Appl
c 396	16	0.5	1801	10	US-09-746-359A-22	Sequence 22, Appl	c 469	16	0.5	2565	9	US-09-754-853A-40	Sequence 40, Appl
c 397	16	0.5	1829	10	US-09-925-302-354	Sequence 354, App	c 470	16	0.5	2565	9	US-09-754-853A-41	Sequence 41, Appl
c 398	16	0.5	1851	9	US-09-989-920-36	Sequence 36, Appl	c 471	16	0.5	2565	9	US-09-754-853A-42	Sequence 42, Appl
c 399	16	0.5	1890	10	US-09-880-107-3858	Sequence 3858, App	c 472	16	0.5	2565	9	US-09-754-853A-43	Sequence 43, Appl
c 400	16	0.5	1891	10	US-09-969-708-146	Sequence 146, App	c 473	16	0.5	2567	10	US-09-175-254-1	Sequence 1, Appli
c 401	16	0.5	1893	10	US-09-815-242-8019	Sequence 8019, App	c 474	16	0.5	2617	9	US-09-974-298-183	Sequence 183, Appl
c 402	16	0.5	1894	9	US-09-464-099A-66	Sequence 66, Appl	c 475	16	0.5	2622	10	US-09-950-370-3	Sequence 3, Appli
c 403	16	0.5	1894	9	US-09-861-696-66	Sequence 66, Appl	c 476	16	0.5	2634	9	US-09-754-853A-6	Sequence 6, Appli
c 404	16	0.5	1929	9	US-10-159-749-7	Sequence 7, Appli	c 477	16	0.5	2634	9	US-09-754-853A-28	Sequence 28, Appl
c 405	16	0.5	1929	10	US-09-969-708-100	Sequence 100, App	c 478	16	0.5	2634	9	US-09-754-853A-29	Sequence 29, Appl
c 406	16	0.5	1929	10	US-09-880-107-1636	Sequence 1636, App	c 479	16	0.5	2634	9	US-09-754-853A-30	Sequence 30, Appl
c 407	16	0.5	1931	10	US-09-925-301-454	Sequence 454, App	c 480	16	0.5	2634	9	US-09-754-853A-31	Sequence 31, Appl
c 408	16	0.5	1932	10	US-09-815-242-4615	Sequence 4615, App	c 481	16	0.5	2634	9	US-09-754-853A-32	Sequence 32, Appl
c 409	16	0.5	1935	10	US-09-815-242-8597	Sequence 8597, App	c 482	16	0.5	2634	9	US-09-754-853A-33	Sequence 33, Appl
c 410	16	0.5	1935	10	US-09-815-242-8900	Sequence 8900, App	c 483	16	0.5	2634	9	US-09-754-853A-34	Sequence 34, Appl
c 411	16	0.5	1951	9	US-09-736-968A-104	Sequence 104, App	c 484	16	0.5	2634	9	US-09-754-853A-35	Sequence 35, Appl
c 412	16	0.5	1956	9	US-09-938-842A-2561	Sequence 2561, App	c 485	16	0.5	2667	10	US-09-810-796-3	Sequence 3, Appli
c 413	16	0.5	1969	10	US-09-925-300-155	Sequence 155, App	c 486	16	0.5	2694	10	US-09-810-796-2	Sequence 2, Appli
c 414	16	0.5	1997	10	US-09-795-006A-21	Sequence 21, Appl	c 487	16	0.5	2701	10	US-09-784-877-3378	Sequence 3378, App
c 415	16	0.5	2000	9	US-09-938-842A-5058	Sequence 5058, App	c 488	16	0.5	2706	10	US-09-950-370-13	Sequence 13, Appl
c 416	16	0.5	2000	9	US-09-938-842A-5088	Sequence 5088, App	c 489	16	0.5	2706	12	US-10-044-090-122	Sequence 122, App
c 417	16	0.5	2000	10	US-09-954-456-1237	Sequence 1237, App	c 490	16	0.5	2720	9	US-10-174-590-137	Sequence 137, App
c 418	16	0.5	2000	10	US-09-997-165-3	Sequence 3, Appli	c 491	16	0.5	2720	9	US-10-176-758-137	Sequence 137, App
c 419	16	0.5	2000	10	US-09-887-576-837	Sequence 837, App	c 492	16	0.5	2720	9	US-10-175-737-137	Sequence 137, App
c 420	16	0.5	2002	10	US-09-925-300-592	Sequence 592, App	c 493	16	0.5	2720	10	US-09-892-5	Sequence 5, Appli
c 421	16	0.5	2007	9	US-09-938-842A-2070	Sequence 2070, App	c 494	16	0.5	2720	12	US-10-052-586-137	Sequence 137, App
c 422	16	0.5	2025	10	US-09-971-309-17	Sequence 17, Appl	c 495	16	0.5	2725	10	US-09-962-436-40	Sequence 40, Appl
c 423	16	0.5	2029	9	US-09-738-626-2417	Sequence 2417, App	c 496	16	0.5	2730	9	US-09-935-371-39	Sequence 39, Appl
c 424	16	0.5	2037	10	US-09-822-830A-25	Sequence 25, Appl	c 497	16	0.5	2733	10	US-09-789-561-59	Sequence 59, Appl
c 425	16	0.5	2056	10	US-09-925-301-482	Sequence 482, App	c 498	16	0.5	2739	10	US-09-864-864-287	Sequence 287, App
c 426	16	0.5	2066	10	US-09-835-996A-14	Sequence 14, Appl	c 499	16	0.5	2739	10	US-09-864-864-322	Sequence 322, App
c 427	16	0.5	2130	9	US-09-740-369-1	Sequence 1, Appli	c 500	16	0.5	2772	10	US-09-825-147-1	Sequence 1, Appli
c 428	16	0.5	2136	9	US-09-981-876-120	Sequence 120, App	c 501	16	0.5	2784	9	US-10-071-766-42	Sequence 42, Appl
c 429	16	0.5	2143	10	US-09-820-893-44	Sequence 44, Appl	c 502	16	0.5	2814	9	US-09-935-371-38	Sequence 38, Appl
c 430	16	0.5	2143	10	US-09-833-381-1022	Sequence 1022, App	c 503	16	0.5	2832	12	US-10-115-178-2	Sequence 2, Appli
c 431	16	0.5	2161	8	US-08-834-666A-17	Sequence 17, Appl	c 504	16	0.5	2838	10	US-09-954-456-700	Sequence 700, App
c 432	16	0.5	2193	8	US-08-910-386A-15	Sequence 15, Appl	c 505	16	0.5	2838	10	US-09-954-456-966	Sequence 966, App
c 433	16	0.5	2196	10	US-09-925-300-504	Sequence 504, App	c 506	16	0.5	2843	9	US-09-954-456-1143	Sequence 1143, App
c 434	16	0.5	2230	10	US-09-925-299-87	Sequence 87, Appl	c 507	16	0.5	2843	9	US-09-854-133-430	Sequence 430, App
c 435	16	0.5	2248	8	US-08-834-666A-15	Sequence 15, Appl	c 508	16	0.5	2843	10	US-09-738-973-430	Sequence 430, App
c 436	16	0.5	2261	10	US-09-940-919-3	Sequence 3, Appli	c 509	16	0.5	2853	10	US-09-905-983-45	Sequence 45, Appl
c 437	16	0.5	2261	10	US-09-917-800A-484	Sequence 484, App	c 510	16	0.5	2883	9	US-10-174-590-157	Sequence 157, App
c 438	16	0.5	2279	10	US-09-954-456-269	Sequence 269, App	c 511	16	0.5	2883	9	US-10-176-758-137	Sequence 137, App
c 439	16	0.5	2279	10	US-09-954-456-950	Sequence 950, App	c 512	16	0.5	2883	9	US-10-175-737-157	Sequence 157, App
c 440	16	0.5	2279	10	US-09-954-456-1600	Sequence 1600, App	c 513	16	0.5	2883	12	US-10-052-586-157	Sequence 157, App
c 441	16	0.5	2289	10	US-09-815-242-7422	Sequence 7422, App	c 514	16	0.5	2910	10	US-09-833-790-418	Sequence 418, App
c 442	16	0.5	2289	10	US-09-850-964-3	Sequence 3, Appli	c 515	16	0.5	2910	10	US-09-880-107-3784	Sequence 3784, App
c 443	16	0.5	2294	10	US-09-735-705-123	Sequence 123, App	c 516	16	0.5	2934	10	US-09-815-242-4326	Sequence 4326, App
c 444	16	0.5	2294	10	US-09-850-716A-123	Sequence 123, App	c 517	16	0.5	2935	9	US-10-047-412A-27	Sequence 27, Appl
c 445	16	0.5	2294	10	US-09-897-778-123	Sequence 123, App	c 518	16	0.5	2936	9	US-10-108-605-54	Sequence 54, Appl
c 446	16	0.5	2298	9	US-09-954-531-1386	Sequence 1386, App	c 519	16	0.5	2956	9	US-09-978-295A-89	Sequence 89, Appl
c 447	16	0.5	2311	10	US-09-925-302-244	Sequence 244, App	c 520	16	0.5	2956	9	US-09-978-697-89	Sequence 89, Appl
c 448	16	0.5	2316	9	US-10-028-072-115	Sequence 115, App	c 521	16	0.5	2956	9	US-09-978-192A-89	Sequence 89, Appl
c 449	16	0.5	2316	9	US-09-738-626-3161	Sequence 3161, App	c 522	16	0.5	2956	9	US-09-999-832A-89	Sequence 89, Appl
c 450	16	0.5	2322	9	US-10-045-815-7	Sequence 7, Appli	c 523	16	0.5	2956	9	US-09-978-189-89	Sequence 89, Appl
c 451	16	0.5	2335	10	US-09-742-684-3	Sequence 3, Appli	c 524	16	0.5	2956	9	US-10-174-590-73	Sequence 73, Appl
c 452	16	0.5	2405	9	US-09-981-353-156	Sequence 156, App	c 525	16	0.5	2956	9	US-10-176-758-73	Sequence 73, Appl
c 453	16	0.5	2416	9	US-10-045-815-5	Sequence 5, Appli	c 526	16	0.5	2956	12	US-10-175-737-73	Sequence 73, Appl
c 454	16	0.5	2418	9	US-09-938-842A-872	Sequence 872, App	c 527	16	0.5	2956	12	US-10-052-586-73	Sequence 73, Appl
c 455	16	0.5	2435	10	US-09-823-038A-40	Sequence 40, Appl	c 528	16	0.5	2964	10	US-09-815-242-8094	Sequence 8094, App
c 456	16	0.5	2453	12	US-10-005-858A-1	Sequence 1, Appl	c 529	16	0.5	2981	10	US-09-950-370-1	Sequence 1, Appli
c 457	16	0.5	2466	12	US-10-044-090-251	Sequence 251, App	c 530	16	0.5	3016	10	US-09-925-301-411	Sequence 411, App



c 531	16	0.5	3024	10	US-09-833-381-889	Sequence 889, App	c 604	16	0.5	6751	10	US-09-761-466-5	Sequence 5, Appl
c 532	16	0.5	3035	10	US-09-864-864-311	Sequence 311, App	c 605	16	0.5	8119	10	US-09-802-669-45	Sequence 45, Appl
c 533	16	0.5	3059	10	US-09-770-595A-2	Sequence 2, Appl	c 606	16	0.5	8372	9	US-09-902-941-1893	Sequence 1893, Ap
c 534	16	0.5	3065	10	US-09-950-370-11	Sequence 11, Appl	c 607	16	0.5	8372	9	US-09-849-626-1893	Sequence 1893, Ap
c 535	16	0.5	3071	10	US-09-810-796-1	Sequence 1, Appl	c 608	16	0.5	9567	12	US-10-016-768-12	Sequence 12, Appl
c 536	16	0.5	3074	10	US-09-813-148-1	Sequence 1, Appl	c 609	16	0.5	9574	12	US-10-016-768-9	Sequence 9, Appl
c 537	16	0.5	3111	10	US-09-825-147-3	Sequence 3, Appl	c 610	16	0.5	9581	10	US-09-764-860-895	Sequence 895, App
c 538	16	0.5	3124	9	US-09-974-298-128	Sequence 128, App	c 611	16	0.5	10883	10	US-09-853-386-13	Sequence 13, Appl
c 539	16	0.5	3191	9	US-09-989-920-54	Sequence 54, Appl	c 612	16	0.5	11186	10	US-09-957-997-1	Sequence 1, Appl
c 540	16	0.5	3222	10	US-09-974-300-2654	Sequence 2654, Ap	c 613	16	0.5	11580	10	US-09-880-107-3436	Sequence 3436, Ap
c 541	16	0.5	3274	10	US-09-070-927A-350	Sequence 350, App	c 614	16	0.5	11945	9	US-09-922-549B-67	Sequence 67, Appl
c 542	16	0.5	3313	9	US-09-978-295A-389	Sequence 389, App	c 615	16	0.5	12606	10	US-09-957-974-2	Sequence 2, Appl
c 543	16	0.5	3313	9	US-09-978-697-389	Sequence 389, App	c 616	16	0.5	12728	10	US-09-922-549B-65	Sequence 65, Appl
c 544	16	0.5	3313	9	US-09-978-192A-389	Sequence 389, App	c 617	16	0.5	13341	8	US-08-910-386A-1	Sequence 1, Appl
c 545	16	0.5	3313	9	US-09-999-832A-389	Sequence 389, App	c 618	16	0.5	14175	9	US-09-764-868-1474	Sequence 1474, Ap
c 546	16	0.5	3313	9	US-09-978-189-389	Sequence 389, App	c 619	16	0.5	15613	10	US-09-764-877-3525	Sequence 3525, Ap
c 547	16	0.5	3313	9	US-10-174-590-77	Sequence 77, Appl	c 620	16	0.5	16545	10	US-09-835-996A-32	Sequence 32, Appl
c 548	16	0.5	3313	9	US-10-176-758-77	Sequence 77, Appl	c 621	16	0.5	17379	10	US-09-764-860-861	Sequence 861, App
c 549	16	0.5	3313	9	US-10-175-737-77	Sequence 77, Appl	c 622	16	0.5	31412	9	US-10-109-551-3	Sequence 3, Appl
c 550	16	0.5	3313	12	US-10-052-586-77	Sequence 77, Appl	c 623	16	0.5	41936	10	US-09-967-768A-116	Sequence 116, App
c 551	16	0.5	3358	9	US-09-954-531-1354	Sequence 1354, Ap	c 624	16	0.5	42999	9	US-09-799-462A-17	Sequence 17, Appl
c 552	16	0.5	3376	10	US-09-834-975-775	Sequence 775, App	c 625	16	0.5	42999	9	US-09-799-462A-17	Sequence 17, Appl
c 553	16	0.5	3400	10	US-09-987-025-1	Sequence 1, Appl	c 626	16	0.5	42999	9	US-10-125-767-17	Sequence 17, Appl
c 554	16	0.5	3447	10	US-09-815-242-9728	Sequence 9728, Ap	c 627	16	0.5	42999	9	US-10-125-767-17	Sequence 17, Appl
c 555	16	0.5	3516	10	US-09-746-359A-10	Sequence 10, Appl	c 628	16	0.5	78056	9	US-10-109-551-1	Sequence 1, Appl
c 556	16	0.5	3609	10	US-09-799-875-6	Sequence 6, Appl	c 629	16	0.5	80959	9	US-09-858-546-3	Sequence 3, Appl
c 557	16	0.5	3637	10	US-09-910-150-1	Sequence 1, Appl	c 630	16	0.5	167343	10	US-09-962-436-281	Sequence 281, App
c 558	16	0.5	3641	9	US-10-078-650-1	Sequence 1, Appl	c 631	16	0.5	167343	10	US-09-964-824A-273	Sequence 273, App
c 559	16	0.5	3657	10	US-09-728-432-4	Sequence 4, Appl	c 632	16	0.5	174424	10	US-09-967-768A-314	Sequence 314, App
c 560	16	0.5	3887	10	US-09-790-318-3	Sequence 3, Appl	c 633	16	0.5	242487	9	US-10-026-188-3	Sequence 3, Appl
c 561	16	0.5	3893	9	US-10-114-893-179	Sequence 179, App	c 634	16	0.5	335913	9	US-09-754-853A-2	Sequence 2, Appl
c 562	16	0.5	3830	9	US-10-078-770-185	Sequence 185, App	c 635	16	0.5	335913	9	US-09-754-853A-3	Sequence 3, Appl
c 563	16	0.5	3920	10	US-09-923-684-1	Sequence 1, Appl	c 636	16	0.5	368004	9	US-09-949-654-3	Sequence 3, Appl
c 564	16	0.5	3954	10	US-09-815-242-4063	Sequence 4063, Ap	c 637	16	0.5	659158	9	US-09-771-208-20	Sequence 20, Appl
c 565	16	0.5	3991	12	US-10-044-090-666	Sequence 666, App	c 638	16	0.5	684973	10	US-09-263-959-1	Sequence 1, Appl
c 566	16	0.5	4098	10	US-09-962-436-37	Sequence 37, Appl	c 639	16	0.5	1503841	9	US-09-946-807-1	Sequence 1, Appl
c 567	16	0.5	4134	9	US-09-975-553-17	Sequence 17, Appl	c 640	16	0.5	1503841	10	US-09-946-807-1	Sequence 1, Appl
c 568	16	0.5	4286	10	US-09-699-634A-1	Sequence 1, Appl	c 641	16	0.5	1503841	10	US-09-795-668-1	Sequence 1, Appl
c 569	16	0.5	4339	12	US-10-002-600-50	Sequence 50, Appl	c 642	16	0.5	1503841	10	US-09-795-668-1	Sequence 1, Appl
c 570	16	0.5	4368	9	US-09-981-353-15	Sequence 15, Appl	c 643	16	0.5	1503841	10	US-09-795-686-1	Sequence 1, Appl
c 571	16	0.5	4501	9	US-09-754-853A-10	Sequence 10, Appl	c 644	16	0.5	1503841	10	US-09-795-686-1	Sequence 1, Appl
c 572	16	0.5	4501	9	US-09-754-853A-11	Sequence 11, Appl	c 645	16	0.5	3309400	9	US-09-738-626-1	Sequence 10, Appl
c 573	16	0.5	4508	9	US-09-754-853A-18	Sequence 18, Appl	c 646	15	0.5	21	10	US-09-828-034-10	Sequence 4, Appl
c 574	16	0.5	4508	9	US-09-754-853A-19	Sequence 19, Appl	c 647	15	0.5	30	9	US-10-011-635A-4	Sequence 16, Appl
c 575	16	0.5	4508	9	US-09-754-853A-20	Sequence 20, Appl	c 648	15	0.5	50	10	US-09-782-650-16	Sequence 17, Appl
c 576	16	0.5	4508	9	US-09-754-853A-21	Sequence 21, Appl	c 649	15	0.5	51	10	US-09-782-650-17	Sequence 17, Appl
c 577	16	0.5	4508	9	US-09-754-853A-22	Sequence 22, Appl	c 650	15	0.5	80	10	US-09-864-761-27458	Sequence 27458, A
c 578	16	0.5	4508	9	US-09-754-853A-23	Sequence 23, Appl	c 651	15	0.5	92	10	US-09-864-761-21237	Sequence 21237, A
c 579	16	0.5	4522	9	US-09-754-853A-12	Sequence 12, Appl	c 652	15	0.5	113	9	US-09-764-868-1360	Sequence 1360, Ap
c 580	16	0.5	4522	9	US-09-754-853A-13	Sequence 13, Appl	c 653	15	0.5	118	10	US-09-770-696-700	Sequence 700, App
c 581	16	0.5	4523	9	US-09-754-853A-8	Sequence 8, Appl	c 654	15	0.5	119	10	US-09-865-499-4	Sequence 4, Appl
c 582	16	0.5	4523	9	US-09-754-853A-9	Sequence 9, Appl	c 655	15	0.5	122	10	US-09-864-761-25491	Sequence 25491, A
c 583	16	0.5	4523	9	US-09-754-853A-14	Sequence 14, Appl	c 656	15	0.5	129	10	US-09-923-876-1595	Sequence 1595, Ap
c 584	16	0.5	4523	9	US-09-754-853A-15	Sequence 15, Appl	c 657	15	0.5	131	10	US-09-998-598-982	Sequence 982, App
c 585	16	0.5	4523	9	US-09-754-853A-16	Sequence 16, Appl	c 658	15	0.5	133	10	US-09-923-876-2863	Sequence 2863, Ap
c 586	16	0.5	4523	9	US-09-754-853A-17	Sequence 17, Appl	c 659	15	0.5	134	10	US-09-983-965-2662	Sequence 2662, Ap
c 587	16	0.5	4600	9	US-09-736-457-1797	Sequence 1797, Ap	c 660	15	0.5	138	10	US-09-244-694-129	Sequence 129, App
c 588	16	0.5	4600	9	US-09-902-941-1797	Sequence 1797, Ap	c 661	15	0.5	143	10	US-09-864-761-26862	Sequence 26862, A
c 589	16	0.5	4600	9	US-09-849-626-1797	Sequence 1797, Ap	c 662	15	0.5	143	10	US-09-974-300-5308	Sequence 5308, Ap
c 590	16	0.5	4739	9	US-09-954-531-586	Sequence 586, App	c 663	15	0.5	147	10	US-09-924-035A-247	Sequence 247, App
c 591	16	0.5	4780	10	US-09-962-436-287	Sequence 287, App	c 664	15	0.5	149	10	US-09-878-574-14501	Sequence 14501, A
c 592	16	0.5	5025	10	US-09-960-253-176	Sequence 176, App	c 665	15	0.5	155	10	US-09-770-696-446	Sequence 446, App
c 593	16	0.5	5222	10	US-09-416-384A-1	Sequence 1, Appl	c 666	15	0.5	159	9	US-09-738-626-3296	Sequence 3296, Ap
c 594	16	0.5	5427	10	US-09-981-353-159	Sequence 159, App	c 667	15	0.5	161	10	US-09-884-441-35	Sequence 35, Appl
c 595	16	0.5	5566	10	US-09-416-384A-4	Sequence 4, Appl	c 668	15	0.5	168	10	US-09-879-536-222	Sequence 222, App
c 596	16	0.5	5597	10	US-09-885-072-4	Sequence 4, Appl	c 669	15	0.5	172	10	US-09-878-574-12516	Sequence 12516, A
c 597	16	0.5	5597	10	US-09-986-552-4	Sequence 4, Appl	c 670	15	0.5	175	10	US-09-954-456-1477	Sequence 1477, Ap
c 598	16	0.5	5756	9	US-09-548-880-1	Sequence 1, Appl	c 671	15	0.5	184	10	US-09-864-761-31978	Sequence 31978, A
c 599	16	0.5	5983	10	US-09-799-875-4	Sequence 4, Appl	c 672	15	0.5	194	10	US-09-864-761-32733	Sequence 32733, A
c 600	16	0.5	6057	12	US-10-044-090-239	Sequence 239, App	c 673	15	0.5	199	10	US-09-864-761-32733	Sequence 22166, A
c 601	16	0.5	6181	9	US-09-764-868-1280	Sequence 1280, Ap	c 674	15	0.5	203	10	US-09-294-093B-4713	Sequence 4713, Ap
c 602	16	0.5	6356	10	US-09-770-595A-1	Sequence 1, Appl	c 675	15	0.5	203	10	US-09-878-574-8445	Sequence 8445, Ap
c 603	16	0.5	6693	9	US-10-071-766-135	Sequence 135, App	c 676	15	0.5	204	12	US-10-010-901-10	Sequence 10, Appl

c 677	15	0.5	207	10	US-09-974-300-4196	Sequence 4196, Ap	750	15	0.5	307	10	US-09-867-701-10620	Sequence 10620, A
c 678	15	0.5	213	10	US-09-923-876-178	Sequence 178, App	751	15	0.5	311	10	US-09-960-352-14967	Sequence 14967, A
c 679	15	0.5	215	10	US-09-815-242-1079	Sequence 1079, Ap	c 752	15	0.5	322	10	US-09-960-352-6429	Sequence 6429, Ap
c 680	15	0.5	216	9	US-09-796-692-7248	Sequence 7248, Ap	753	15	0.5	327	10	US-09-880-107-783	Sequence 783, App
c 681	15	0.5	217	10	US-09-878-574-12062	Sequence 12062, A	c 754	15	0.5	330	10	US-09-815-242-2247	Sequence 2247, Ap
c 682	15	0.5	217	10	US-09-962-832-74	Sequence 74, Appl	755	15	0.5	332	10	US-09-294-093B-2203	Sequence 2203, Ap
c 683	15	0.5	222	10	US-09-974-300-41139	Sequence 4139, App	756	15	0.5	337	9	US-10-015-219-330	Sequence 330, App
c 684	15	0.5	226	10	US-09-728-445-159	Sequence 159, App	757	15	0.5	337	9	US-09-777-564-330	Sequence 330, App
c 685	15	0.5	228	10	US-09-815-242-819	Sequence 819, App	758	15	0.5	338	9	US-09-933-797-130	Sequence 130, App
c 686	15	0.5	231	10	US-09-878-574-8627	Sequence 8627, Ap	c 759	15	0.5	339	9	US-09-930-455-161	Sequence 161, App
c 687	15	0.5	231	10	US-09-860-107-1352	Sequence 1352, Ap	c 760	15	0.5	342	10	US-09-960-352-1348	Sequence 1348, Ap
c 688	15	0.5	233	10	US-09-960-352-6124	Sequence 6124, Ap	c 761	15	0.5	342	10	US-09-960-352-13723	Sequence 13723, A
c 689	15	0.5	238	10	US-09-983-965-3902	Sequence 3902, Ap	c 762	15	0.5	343	10	US-09-960-352-14433	Sequence 14433, A
c 690	15	0.5	246	10	US-09-878-574-6208	Sequence 6208, Ap	763	15	0.5	345	10	US-09-878-574-863	Sequence 863, App
c 691	15	0.5	247	10	US-09-923-876-2971	Sequence 2971, Ap	764	15	0.5	346	9	US-09-925-664-50	Sequence 50, Appl
c 692	15	0.5	249	10	US-09-783-590-10662	Sequence 10662, A	c 765	15	0.5	348	10	US-09-960-352-13719	Sequence 13719, A
c 693	15	0.5	249	10	US-09-883-825-40	Sequence 40, Appl	766	15	0.5	349	9	US-09-986-140-94	Sequence 94, Appl
c 694	15	0.5	251	9	US-09-933-797-319	Sequence 319, App	c 767	15	0.5	350	10	US-09-983-965-840	Sequence 840, App
c 695	15	0.5	251	10	US-09-878-574-7980	Sequence 7980, Ap	c 768	15	0.5	351	10	US-09-867-701-4299	Sequence 4299, Ap
c 696	15	0.5	252	10	US-09-923-876-496	Sequence 496, App	769	15	0.5	353	10	US-09-960-352-12027	Sequence 12027, A
c 697	15	0.5	252	10	US-09-878-574-15770	Sequence 15770, A	770	15	0.5	355	10	US-09-867-701-4602	Sequence 4602, Ap
c 698	15	0.5	254	10	US-09-878-574-10714	Sequence 10714, A	771	15	0.5	355	10	US-09-244-694-138	Sequence 138, App
c 699	15	0.5	254	10	US-09-878-574-11391	Sequence 11391, A	c 772	15	0.5	363	10	US-09-878-574-391	Sequence 391, App
c 700	15	0.5	255	9	US-09-738-626-666	Sequence 666, App	c 773	15	0.5	366	10	US-09-960-352-3994	Sequence 3994, Ap
c 701	15	0.5	258	10	US-09-923-876-1914	Sequence 1914, Ap	c 774	15	0.5	367	10	US-09-867-701-8008	Sequence 8008, Ap
c 702	15	0.5	258	10	US-09-878-574-10667	Sequence 10667, A	775	15	0.5	369	10	US-09-960-352-10070	Sequence 10070, A
c 703	15	0.5	259	10	US-09-923-876-1666	Sequence 1666, Ap	776	15	0.5	369	10	US-09-983-965-3566	Sequence 3566, Ap
c 704	15	0.5	259	10	US-09-912-020-471	Sequence 471, App	777	15	0.5	371	10	US-09-960-352-8510	Sequence 8510, Ap
c 705	15	0.5	260	10	US-09-912-020-67	Sequence 67, Appl	c 778	15	0.5	373	10	US-09-878-574-2299	Sequence 2299, Ap
c 706	15	0.5	260	10	US-09-878-574-11123	Sequence 11123, A	c 779	15	0.5	375	10	US-09-867-701-3254	Sequence 3254, Ap
c 707	15	0.5	260	10	US-09-878-574-13054	Sequence 13054, A	780	15	0.5	376	9	US-09-796-692-3399	Sequence 3399, Ap
c 708	15	0.5	261	10	US-09-923-876-695	Sequence 695, App	781	15	0.5	378	10	US-09-878-574-1986	Sequence 1986, Ap
c 709	15	0.5	261	10	US-09-864-761-17969	Sequence 17969, A	782	15	0.5	381	10	US-09-770-791-173	Sequence 173, App
c 710	15	0.5	262	10	US-09-923-876-3431	Sequence 3431, Ap	783	15	0.5	381	10	US-09-983-965-5851	Sequence 5851, Ap
c 711	15	0.5	264	10	US-09-294-093B-730	Sequence 730, App	c 784	15	0.5	381	10	US-09-999-256-12	Sequence 12, Appl
c 712	15	0.5	265	9	US-10-040-739-1219	Sequence 1219, App	785	15	0.5	382	10	US-09-880-107-2794	Sequence 2794, Ap
c 713	15	0.5	265	10	US-09-294-093B-792	Sequence 792, App	c 786	15	0.5	383	10	US-09-867-701-5594	Sequence 5594, Ap
c 714	15	0.5	265	10	US-09-878-574-9152	Sequence 9152, Ap	c 787	15	0.5	384	10	US-09-960-352-8976	Sequence 8976, Ap
c 715	15	0.5	267	10	US-09-923-876-5854	Sequence 5854, Ap	c 788	15	0.5	386	10	US-09-960-352-11656	Sequence 11656, A
c 716	15	0.5	268	10	US-09-880-107-2915	Sequence 2915, Ap	c 789	15	0.5	388	10	US-09-347-331-5	Sequence 5, Appl
c 717	15	0.5	271	10	US-09-923-876-5076	Sequence 5076, Ap	c 790	15	0.5	388	10	US-09-878-574-528	Sequence 528, App
c 718	15	0.5	271	10	US-09-878-574-7633	Sequence 7633, Ap	c 791	15	0.5	388	10	US-09-867-701-4534	Sequence 4534, Ap
c 719	15	0.5	273	10	US-09-878-574-8991	Sequence 8991, Ap	792	15	0.5	391	10	US-09-770-791-9	Sequence 9, Appl
c 720	15	0.5	273	10	US-09-878-574-15635	Sequence 15635, A	793	15	0.5	392	10	US-09-880-107-353	Sequence 353, App
c 721	15	0.5	274	10	US-09-878-574-9369	Sequence 9369, Ap	794	15	0.5	393	12	US-09-920-300A-1581	Sequence 1581, Ap
c 722	15	0.5	276	10	US-09-793-306-151	Sequence 151, App	795	15	0.5	393	10	US-10-033-528-1581	Sequence 1581, Ap
c 723	15	0.5	276	10	US-09-960-352-11026	Sequence 11026, A	c 796	15	0.5	396	10	US-09-864-761-5393	Sequence 5393, Ap
c 724	15	0.5	280	10	US-09-923-876-6034	Sequence 6034, Ap	797	15	0.5	396	10	US-09-960-352-870	Sequence 870, App
c 725	15	0.5	280	10	US-09-878-574-8050	Sequence 8050, Ap	798	15	0.5	397	10	US-09-974-300-3551	Sequence 3551, Ap
c 726	15	0.5	281	10	US-09-294-093B-5703	Sequence 5703, Ap	799	15	0.5	398	10	US-09-960-352-12334	Sequence 12334, A
c 727	15	0.5	281	10	US-09-878-574-15012	Sequence 15012, A	800	15	0.5	398	10	US-09-983-965-184	Sequence 184, App
c 728	15	0.5	283	10	US-09-294-093B-931	Sequence 931, App	801	15	0.5	401	10	US-09-920-345-24	Sequence 24, Appl
c 729	15	0.5	283	10	US-09-878-574-5469	Sequence 5469, Ap	802	15	0.5	402	10	US-09-974-300-3122	Sequence 3122, Ap
c 730	15	0.5	284	10	US-09-867-701-2907	Sequence 2907, Ap	c 803	15	0.5	404	10	US-09-878-574-3908	Sequence 3908, Ap
c 731	15	0.5	285	10	US-09-912-020-191	Sequence 191, App	c 804	15	0.5	405	10	US-09-983-965-5418	Sequence 5418, Ap
c 732	15	0.5	288	10	US-09-783-590-4734	Sequence 4734, Ap	805	15	0.5	406	10	US-09-783-590-11317	Sequence 11317, A
c 733	15	0.5	288	10	US-09-867-701-5557	Sequence 5557, Ap	806	15	0.5	406	10	US-09-960-352-1716	Sequence 1716, Ap
c 734	15	0.5	289	10	US-09-294-093B-5584	Sequence 5584, Ap	c 807	15	0.5	406	10	US-09-949-192-20	Sequence 20, Appl
c 735	15	0.5	290	10	US-09-294-093B-243	Sequence 243, App	c 808	15	0.5	407	9	US-09-933-797-759	Sequence 759, App
c 736	15	0.5	294	10	US-09-244-694-137	Sequence 137, App	809	15	0.5	407	9	US-10-025-380-174	Sequence 174, App
c 737	15	0.5	294	10	US-09-864-761-19936	Sequence 19936, A	810	15	0.5	407	10	US-09-922-217-174	Sequence 174, App
c 738	15	0.5	294	10	US-09-783-590-4	Sequence 4, Appl	811	15	0.5	407	10	US-09-833-263-174	Sequence 174, App
c 739	15	0.5	296	10	US-09-294-093B-6090	Sequence 6090, Ap	c 812	15	0.5	408	10	US-09-878-574-1046	Sequence 1046, Ap
c 740	15	0.5	297	10	US-09-294-093B-1856	Sequence 1856, Ap	813	15	0.5	410	10	US-09-783-590-7404	Sequence 7404, Ap
c 741	15	0.5	300	10	US-09-294-093B-3923	Sequence 3923, Ap	814	15	0.5	410	10	US-09-960-352-8643	Sequence 8643, Ap
c 742	15	0.5	300	10	US-09-941-997-1	Sequence 1, Appl	c 815	15	0.5	411	9	US-10-040-739-599	Sequence 599, App
c 743	15	0.5	302	9	US-09-796-692-6537	Sequence 6537, Ap	816	15	0.5	411	9	US-09-854-133-361	Sequence 361, App
c 744	15	0.5	302	10	US-09-867-701-7885	Sequence 7885, Ap	817	15	0.5	411	10	US-09-738-973-361	Sequence 8231, Ap
c 745	15	0.5	305	9	US-09-928-457-69	Sequence 69, Appl	c 818	15	0.5	412	10	US-09-960-352-8231	Sequence 8231, Ap
c 746	15	0.5	305	9	US-10-040-739-844	Sequence 844, App	c 819	15	0.5	413	10	US-09-919-580-301	Sequence 301, App
c 747	15	0.5	306	10	US-09-815-242-1364	Sequence 1364, Ap	c 820	15	0.5	413	10	US-09-960-352-138	Sequence 138, App
c 748	15	0.5	306	10	US-09-815-242-1523	Sequence 1523, Ap	c 821	15	0.5	414	10	US-09-960-352-1603	Sequence 1603, Ap
c 749	15	0.5	306	10	US-09-815-242-1649	Sequence 1649, Ap	822	15	0.5	414	10	US-09-815-242-1060	Sequence 1060, Ap

c 823	15	0.5	414	10	US-09-960-352-13586	Sequence 13586, A	c 896	15	0.5	479	10	US-09-864-761-1205	Sequence 1205, Ap
c 824	15	0.5	415	10	US-09-783-590-7545	Sequence 7545, Ap	c 897	15	0.5	479	10	US-09-864-761-10224	Sequence 10224, A
c 825	15	0.5	416	10	US-09-960-352-7067	Sequence 7067, Ap	c 898	15	0.5	480	10	US-09-864-761-133	Sequence 133, Ap
c 826	15	0.5	417	10	US-09-867-550-731	Sequence 731, App	c 899	15	0.5	481	10	US-09-864-761-13122	Sequence 15122, A
c 827	15	0.5	418	10	US-09-878-574-4107	Sequence 4107, Ap	c 900	15	0.5	484	9	US-09-796-692-9286	Sequence 9286, Ap
c 828	15	0.5	419	10	US-09-960-352-3158	Sequence 3158, Ap	c 901	15	0.5	484	10	US-09-560-863-982	Sequence 982, App
c 829	15	0.5	420	10	US-09-983-965-4281	Sequence 4281, Ap	c 902	15	0.5	486	10	US-09-864-761-14504	Sequence 14504, A
c 830	15	0.5	421	10	US-09-864-761-10577	Sequence 10577, A	c 903	15	0.5	486	10	US-09-682-706-4	Sequence 4, Appl
c 831	15	0.5	422	10	US-09-974-300-7159	Sequence 7159, Ap	c 904	15	0.5	486	10	US-09-917-800A-1171	Sequence 1171, Ap
c 832	15	0.5	423	9	US-09-938-842A-3801	Sequence 3801, Ap	c 905	15	0.5	489	10	US-09-244-694-153	Sequence 153, App
c 833	15	0.5	424	10	US-09-960-352-14599	Sequence 14599, A	c 906	15	0.5	498	10	US-09-783-590-7412	Sequence 7412, Ap
c 834	15	0.5	425	9	US-10-114-893-91	Sequence 91, Appl	c 907	15	0.5	498	10	US-09-778-844-188	Sequence 188, App
c 835	15	0.5	426	10	US-09-876-889-249	Sequence 249, App	c 908	15	0.5	500	10	US-09-917-800A-1273	Sequence 1273, Ap
c 836	15	0.5	427	10	US-09-764-877-152	Sequence 152, App	c 909	15	0.5	501	10	US-09-864-761-8824	Sequence 8824, Ap
c 837	15	0.5	428	10	US-09-960-352-6081	Sequence 6081, Ap	c 910	15	0.5	501	10	US-09-864-761-16425	Sequence 16425, A
c 838	15	0.5	429	10	US-09-878-574-3832	Sequence 3832, Ap	c 911	15	0.5	501	10	US-09-884-441-118	Sequence 118, App
c 839	15	0.5	430	10	US-09-960-352-12560	Sequence 12560, A	c 912	15	0.5	502	10	US-09-833-381-133	Sequence 133, App
c 840	15	0.5	431	10	US-09-983-965-2989	Sequence 2989, Ap	c 913	15	0.5	505	10	US-09-783-590-5959	Sequence 5959, Ap
c 841	15	0.5	432	9	US-10-136-224-11	Sequence 11, Appl	c 914	15	0.5	506	10	US-09-895-828-75	Sequence 12385, A
c 842	15	0.5	433	10	US-09-960-352-7845	Sequence 7845, Ap	c 915	15	0.5	508	10	US-09-783-590-3069	Sequence 75, Appl
c 843	15	0.5	434	10	US-09-983-965-3790	Sequence 3790, Ap	c 916	15	0.5	510	10	US-09-432-239-9	Sequence 3069, Ap
c 844	15	0.5	435	10	US-09-867-701-7104	Sequence 7104, Ap	c 917	15	0.5	510	10	US-09-974-300-809	Sequence 9, Appl
c 845	15	0.5	436	10	US-09-960-352-14799	Sequence 14799, A	c 918	15	0.5	511	10	US-09-783-590-5135	Sequence 809, App
c 846	15	0.5	437	10	US-09-880-107-213	Sequence 213, App	c 919	15	0.5	513	10	US-09-783-590-2724	Sequence 5135, Ap
c 847	15	0.5	438	10	US-09-728-445-698	Sequence 698, App	c 920	15	0.5	515	9	US-09-736-457-151	Sequence 2724, Ap
c 848	15	0.5	439	10	US-09-983-965-498	Sequence 498, App	c 921	15	0.5	515	9	US-09-902-941-151	Sequence 151, App
c 849	15	0.5	440	10	US-09-960-352-9523	Sequence 9523, Ap	c 922	15	0.5	515	9	US-09-849-626-151	Sequence 151, App
c 850	15	0.5	441	10	US-09-983-965-2960	Sequence 2960, Ap	c 923	15	0.5	517	10	US-09-998-598-1387	Sequence 1387, Ap
c 851	15	0.5	442	9	US-09-960-352-11772	Sequence 11772, A	c 924	15	0.5	518	9	US-10-046-935-1816	Sequence 1816, Ap
c 852	15	0.5	443	9	US-09-954-531-914	Sequence 914, App	c 925	15	0.5	518	9	US-09-878-178-1816	Sequence 1816, Ap
c 853	15	0.5	444	9	US-09-954-531-1324	Sequence 1324, Ap	c 926	15	0.5	518	9	US-09-998-598-692	Sequence 692, App
c 854	15	0.5	445	10	US-09-764-868-85	Sequence 85, Appl	c 927	15	0.5	519	10	US-09-864-761-15890	Sequence 15890, A
c 855	15	0.5	446	10	US-09-770-444-835	Sequence 835, App	c 928	15	0.5	520	9	US-09-933-797-8	Sequence 8, Appl
c 856	15	0.5	447	10	US-09-864-761-3157	Sequence 3157, Ap	c 929	15	0.5	524	10	US-09-864-761-15293	Sequence 15293, A
c 857	15	0.5	448	10	US-09-974-300-3588	Sequence 3588, Ap	c 930	15	0.5	527	9	US-09-925-297-272	Sequence 272, App
c 858	15	0.5	449	10	US-09-954-456-221	Sequence 221, App	c 931	15	0.5	529	9	US-09-854-133-682	Sequence 682, App
c 859	15	0.5	450	10	US-09-960-352-1362	Sequence 1362, Ap	c 932	15	0.5	530	9	US-10-073-353-4	Sequence 4, Appl
c 860	15	0.5	451	9	US-09-954-456-1362	Sequence 423, App	c 933	15	0.5	530	10	US-09-867-550-1383	Sequence 1383, Ap
c 861	15	0.5	452	9	US-09-777-564-423	Sequence 423, App	c 934	15	0.5	531	9	US-09-894-844-91	Sequence 91, Appl
c 862	15	0.5	453	9	US-10-015-219-423	Sequence 139, App	c 935	15	0.5	542	9	US-09-796-692-4370	Sequence 4370, Ap
c 863	15	0.5	454	10	US-09-777-564-139	Sequence 139, App	c 936	15	0.5	542	10	US-09-864-761-8238	Sequence 8238, Ap
c 864	15	0.5	455	9	US-09-864-761-2055	Sequence 2055, Ap	c 937	15	0.5	542	10	US-09-864-761-4231	Sequence 4231, Ap
c 865	15	0.5	456	10	US-09-854-133-145	Sequence 145, App	c 938	15	0.5	543	10	US-09-887-576-786	Sequence 786, App
c 866	15	0.5	457	10	US-09-770-444-704	Sequence 704, App	c 939	15	0.5	543	10	US-09-974-300-5138	Sequence 5138, Ap
c 867	15	0.5	458	10	US-09-738-973-145	Sequence 145, App	c 940	15	0.5	545	10	US-09-925-299-382	Sequence 382, App
c 868	15	0.5	459	10	US-09-960-352-5068	Sequence 5068, Ap	c 941	15	0.5	546	10	US-09-815-242-4036	Sequence 4036, Ap
c 869	15	0.5	460	9	US-09-796-692-3075	Sequence 3075, Ap	c 942	15	0.5	550	10	US-09-815-343-1010	Sequence 1010, Ap
c 870	15	0.5	461	10	US-09-867-550-1673	Sequence 1673, Ap	c 943	15	0.5	552	9	US-09-910-664-10	Sequence 10, Appl
c 871	15	0.5	462	9	US-09-796-692-1673	Sequence 4129, Ap	c 944	15	0.5	554	10	US-09-917-800A-1602	Sequence 1602, Ap
c 872	15	0.5	463	10	US-09-770-444-645	Sequence 645, App	c 945	15	0.5	557	9	US-09-828-644-7	Sequence 7, Appl
c 873	15	0.5	464	10	US-09-864-761-4489	Sequence 4489, Ap	c 946	15	0.5	559	10	US-09-810-264-42	Sequence 66, Appl
c 874	15	0.5	465	10	US-09-864-761-10816	Sequence 10816, A	c 947	15	0.5	560	9	US-10-046-935-208	Sequence 42, Appl
c 875	15	0.5	466	10	US-09-867-701-1871	Sequence 1871, Ap	c 948	15	0.5	560	9	US-09-878-178-208	Sequence 208, App
c 876	15	0.5	467	10	US-09-960-352-7568	Sequence 7568, Ap	c 949	15	0.5	562	9	US-09-867-701-5942	Sequence 5942, Ap
c 877	15	0.5	468	10	US-09-783-590-3028	Sequence 3028, Ap	c 950	15	0.5	562	10	US-09-854-133-463	Sequence 463, App
c 878	15	0.5	469	9	US-09-764-868-504	Sequence 504, App	c 951	15	0.5	562	10	US-09-864-761-9787	Sequence 9787, Ap
c 879	15	0.5	470	9	US-09-920-455-171	Sequence 171, App	c 952	15	0.5	562	10	US-09-864-761-16224	Sequence 16224, A
c 880	15	0.5	471	12	US-10-027-559-1	Sequence 1, Appl	c 953	15	0.5	562	10	US-09-738-973-463	Sequence 463, App
c 881	15	0.5	472	10	US-09-833-790-185	Sequence 185, App	c 954	15	0.5	566	10	US-10-062-254-149	Sequence 149, App
c 882	15	0.5	473	9	US-09-954-531-868	Sequence 868, App	c 955	15	0.5	566	10	US-09-925-300-867	Sequence 867, App
c 883	15	0.5	474	10	US-09-954-456-911	Sequence 911, App	c 956	15	0.5	572	10	US-09-864-761-12790	Sequence 12790, A
c 884	15	0.5	475	10	US-09-960-352-1254	Sequence 1254, Ap	c 957	15	0.5	573	9	US-09-796-692-8943	Sequence 8943, Ap
c 885	15	0.5	476	10	US-09-925-300-659	Sequence 659, App	c 958	15	0.5	576	9	US-09-738-626-3297	Sequence 3297, Ap
c 886	15	0.5	477	9	US-09-938-842A-276	Sequence 276, App	c 959	15	0.5	577	10	US-09-864-761-7498	Sequence 7498, Ap
c 887	15	0.5	478	10	US-09-864-761-3022	Sequence 3022, Ap	c 960	15	0.5	582	10	US-09-764-847-47	Sequence 47, Appl
c 888	15	0.5	479	10	US-09-770-444-219	Sequence 219, App	c 961	15	0.5	584	10	US-09-995-494-36	Sequence 36, Appl
c 889	15	0.5	480	10	US-09-867-701-2855	Sequence 2855, Ap	c 962	15	0.5	586	10	US-09-864-761-9940	Sequence 9940, Ap
c 890	15	0.5	481	10	US-09-867-701-6813	Sequence 6813, Ap	c 963	15	0.5	589	10	US-09-917-800A-332	Sequence 332, App
c 891	15	0.5	482	10	US-09-864-761-3999	Sequence 3999, Ap	c 964	15	0.5	594	9	US-09-933-797-745	Sequence 745, App
c 892	15	0.5	483	10	US-09-917-800A-662	Sequence 662, App	c 965	15	0.5				
c 893	15	0.5	484	10	US-09-790-099-7	Sequence 7, Appl	c 966	15	0.5				
c 894	15	0.5	485	10	US-09-864-761-15464	Sequence 15464, A	c 967	15	0.5				
c 895	15	0.5	486	10	US-09-934-249-15	Sequence 15, Appl	c 968	15	0.5				

c 969 15 0.5 597 10 US-09-912-020-200 Sequence 200, App  
c 970 15 0.5 600 10 US-09-770-149-972 Sequence 972, App  
c 971 15 0.5 603 9 US-09-764-904-72 Sequence 72, Appl  
c 972 15 0.5 603 10 US-09-764-860-600 Sequence 600, App  
c 973 15 0.5 605 9 US-10-046-935-1139 Sequence 1139, Ap  
c 974 15 0.5 605 9 US-09-878-178-1139 Sequence 1139, Ap  
c 975 15 0.5 608 10 US-09-813-358-167 Sequence 167, App  
c 976 15 0.5 621 9 US-09-950-933A-5 Sequence 5, Appli  
c 977 15 0.5 621 10 US-09-912-020-134 Sequence 134, App  
c 978 15 0.5 625 9 US-09-854-133-628 Sequence 628, App  
c 979 15 0.5 625 10 US-09-833-790-298 Sequence 298, App  
c 980 15 0.5 631 10 US-09-770-149-684 Sequence 684, App  
c 981 15 0.5 632 10 US-09-770-149-678 Sequence 678, App  
c 982 15 0.5 637 10 US-09-878-574-4574 Sequence 4574, Ap  
c 983 15 0.5 645 12 US-10-036-869-22 Sequence 22, Appl  
c 984 15 0.5 650 12 US-10-078-929-185 Sequence 185, App  
c 985 15 0.5 654 10 US-09-995-598-19 Sequence 19, Appl  
c 986 15 0.5 660 9 US-10-055-364-47 Sequence 47, Appl  
c 987 15 0.5 660 9 US-09-950-933A-18 Sequence 18, Appl  
c 988 15 0.5 660 10 US-09-764-847-1674 Sequence 1674, Ap  
c 989 15 0.5 672 9 US-09-854-133-462 Sequence 462, App  
c 990 15 0.5 672 10 US-09-738-973-462 Sequence 218, App  
c 991 15 0.5 687 10 US-09-778-320-218 Sequence 30, Appl  
c 992 15 0.5 687 10 US-09-770-149-313 Sequence 313, App  
c 993 15 0.5 687 10 US-09-910-689-218 Sequence 218, App  
c 994 15 0.5 687 10 US-09-765-231A-37 Sequence 37, Appl  
c 995 15 0.5 687 12 US-10-010-742-218 Sequence 218, App  
c 996 15 0.5 688 9 US-09-764-868-218 Sequence 218, App  
c 997 15 0.5 693 9 US-09-974-300-1576 Sequence 1576, Ap  
c 998 15 0.5 695 10 US-09-910-943-291 Sequence 291, App  
c 999 15 0.5 707 10 US-09-917-800A-1530 Sequence 1530, Ap  
1000 15 0.5

ALIGNMENTS

RESULT 1  
US-09-928-457-95  
; Sequence 95, Application US/09928457  
; Patent No. US20020164603A1  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: DNA, specific proteins and peptides  
; TITLE OF INVENTION: of the Neisseria meningitidis species bacteria, method  
; TITLE OF INVENTION: for obtaining them and their biological application.  
; NUMBER OF SEQUENCES: 99  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (ORB)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/928,457  
; FILING DATE: 2001-08-14  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/214,759  
; FILING DATE: 199-12-10  
; INFORMATION FOR SEQ ID NO: 95:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 286 base pairs  
; TYPE: nucleotide  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
US-09-928-457-95

Query Match 5.5%; Score 177; DB 9; Length 286;  
Best Local Similarity 100.0%; Pred. No. 3.8e-83;  
Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 865 AATTCGGAGGACAGTAGTACCGCCAAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGAC 924  
Db 1 AATTCGGAGGACAGTAGTACCGCCAAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGAC 60  
Qy 925 GAGGGTATCCGCTGATGCAACAGAGCGGATTACGGCAACTTGTCTTACCACATCGGTAAT 984  
Db 61 GAGGGTATCCGCTGATGCAACAGAGCGGATTACGGCAACTTGTCTTACCACATCGGTAAT 120  
Qy 985 AAAAACATGCTTTTCATTTTTCGCAAGCAATGAGGCACAAAGCTCAGCCCCAACACA 1041  
Db 121 AAAAACATGCTTTTCATTTTTCGCAAGCAATGAGGCACAAAGCTCAGCCCCAACACA 177

RESULT 2  
US-09-867-701-2325  
; Sequence 2325, Application US/09867701  
; Patent No. US20020132237A1  
; GENERAL INFORMATION:  
; APPLICANT: Aglate, Paul A.  
; APPLICANT: Jones, Robert  
; APPLICANT: Harlocker, Susan L.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER  
; FILE REFERENCE: 210121.497  
; CURRENT APPLICATION NUMBER: US/09/867,701  
; CURRENT FILING DATE: 2001-05-29  
; NUMBER OF SEQ ID NOS: 10912  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2325  
; LENGTH: 284  
; TYPE: DNA  
; ORGANISM: Homo sapien  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(284)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-867-701-2325

Query Match 0.6%; Score 19; DB 10; Length 284;  
Best Local Similarity 100.0%; Pred. No. 8;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1263 GGCTCTGCTGCTGCAGAAA 1281  
Db 185 GGCTCTGCTGCTGCAGAAA 203

RESULT 3  
US-09-925-299-125  
; Sequence 125, Application US/09925299  
; Patent No. US20020055627A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA102  
; CURRENT APPLICATION NUMBER: US/09/925,299  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05883  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 1556  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 125  
; LENGTH: 1267  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-299-125

Query Match 0.6%; Score 19; DB 10; Length 1267;  
Best Local Similarity 100.0%; Pred. No. 8.5;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 TAGCGCGCGCGCGGCAC 104  
|||||  
DB 588 TAGCGCGCGCGCGGCAC 606

RESULT 4  
US-09-880-107-2355/c  
; Sequence 2355, Application US/09880107  
; Patent No. US20020142981A1  
; GENERAL INFORMATION:  
; APPLICANT: Horne, Darci T.  
; APPLICANT: Vockley, Joseph G.  
; APPLICANT: Scherif, Uwe  
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer  
; FILE REFERENCE: 44921-5028-WO  
; CURRENT APPLICATION NUMBER: US/09/880,107  
; CURRENT FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: US 60/211,379  
; PRIOR FILING DATE: 2000-06-14  
; PRIOR APPLICATION NUMBER: US 60/237,054  
; PRIOR FILING DATE: 2000-10-02  
; NUMBER OF SEQ ID NOS: 3950  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2355  
; LENGTH: 1551  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 M55671  
US-09-880-107-2355

Query Match 0.6%; Score 19; DB 10; Length 1551;  
Best Local Similarity 100.0%; Pred. No. 8.5;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1258 ACGGCGCTCTGCTGTCG 1276  
|||||  
DB 1430 ACGGCGCTCTGCTGTCG 1412

RESULT 5  
US-09-925-301-306  
; Sequence 306, Application US/09925301  
; Patent No. US20020052308A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA106  
; CURRENT APPLICATION NUMBER: US/09/925,301  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05882  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 1694  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 306  
; LENGTH: 1987  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (731)  
; OTHER INFORMATION: n equals a,t,g, or c  
US-09-925-301-306

Query Match 0.6%; Score 19; DB 10; Length 1987;  
Best Local Similarity 100.0%; Pred. No. 8.6;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1263 GGCTCTGCTGCTGCAGAA 1281  
|||||

DB 1235 GGCTCTGCTGCTGCAGAA 1253

RESULT 6  
US-09-974-300-4730  
; Sequence 4730, Application US/09974300  
; Patent No. US20020146721A1  
; GENERAL INFORMATION:  
; APPLICANT: Berka, Randy M.  
; APPLICANT: Clausen, Ib Groth  
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene  
; TITLE OF INVENTION: Expression  
; FILE REFERENCE: 10085-500-US  
; CURRENT APPLICATION NUMBER: US/09/974,300  
; CURRENT FILING DATE: 2001-10-05  
; PRIOR APPLICATION NUMBER: 09/680,598  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/279,526  
; PRIOR FILING DATE: 2001-03-27  
; NUMBER OF SEQ ID NOS: 8481  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4730  
; LENGTH: 213  
; TYPE: DNA  
; ORGANISM: Bacillus clausii  
US-09-974-300-4730

Query Match 0.6%; Score 18; DB 10; Length 213;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1179 TGAAGCAAGCGTCGTTT 1196  
|||||  
DB 163 TGAAGCAAGCGTCGTTT 180

RESULT 7  
US-09-878-574-15701  
; Sequence 15701, Application US/09878574  
; Patent No. US20020110548A1  
; GENERAL INFORMATION:  
; APPLICANT: Byrum, Joseph R.  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Thompson, Michael D.  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(15401)B  
; CURRENT APPLICATION NUMBER: US/09/878,574  
; CURRENT FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 09/333,535  
; PRIOR FILING DATE: 1999-06-14  
; NUMBER OF SEQ ID NOS: 15775  
; SEQ ID NO 15701  
; LENGTH: 276  
; TYPE: DNA  
; ORGANISM: Glycine max  
; OTHER INFORMATION: Clone ID: 701070414H1  
US-09-878-574-15701

Query Match 0.6%; Score 18; DB 10; Length 276;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 CAGCGAAATCAGCAGCAG 181  
|||||  
DB 85 CAGCGAAATCAGCAGCAG 102

RESULT 8  
US-09-728-445-822/c  
; Sequence 822, Application US/09728445  
; Patent No. US20020102543A1  
; GENERAL INFORMATION:

```
; APPLICANT: Friedrich, Glenn
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. US20020102543A1el Mutated Mammalian Cells and
; TITLE OF INVENTION: Animals
; FILE REFERENCE: LEX-0102-USA
; CURRENT APPLICATION NUMBER: US/09/728,445
; CURRENT FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/168,358
; PRIOR FILING DATE: 1999-12-01
; NUMBER OF SEQ ID NOS: 891
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 822
; LENGTH: 495
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(495)
; OTHER INFORMATION: n = A,T,C or G
US-09-728-445-822

Query Match          0.6%; Score 18; DB 10; Length 495;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 427 GGCAGCATATCCTTTCCC 444
Db 88 GGCAGCATATCCTTTCCC 71

RESULT 9
US-09-974-300-2315/c
; Sequence 2315, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; TITLE OF INVENTION: Methods for Monitoring Multiple Gene
; TITLE OF INVENTION: Expression
; FILE REFERENCE: 10085-500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2315
; LENGTH: 1044
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
US-09-974-300-2315

Query Match          0.6%; Score 18; DB 10; Length 1044;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 845 CCGACCATTTCCAAATAG 862
Db 205 CCGACCATTTCCAAATAG 188

RESULT 10
US-10-098-841-169
; Sequence 169, Application US/10098841
; Publication No. US20020197679A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Xu, Chongjun
```

```
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhao, Qing A.
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wehrman, Tom
; APPLICANT: Zhang, Jie
; APPLICANT: Qian, Xiaohong B.
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20020197679A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2
; CURRENT APPLICATION NUMBER: US/10/098,841
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 331
; SOFTWARE: PT_FL_genes Version 1.0
; SEQ ID NO 169
; LENGTH: 1169
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (114)...(887)
US-10-098-841-169

Query Match          0.6%; Score 18; DB 9; Length 1169;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1063 GAAAAGATGCTCAAAA 1080
Db 602 GAAAAGATGCTCAAAA 619

RESULT 11
US-10-071-766-34
; Sequence 34, Application US/10071766
; Publication No. US20020192678A1
; GENERAL INFORMATION:
; APPLICANT: Huel-Mei Chen
; TITLE OF INVENTION: GENES EXPRESSED IN SENESENCE
; FILE REFERENCE: PA-0043 US
; CURRENT APPLICATION NUMBER: US/10/071,766
; CURRENT FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 144
; SOFTWARE: PERL Program
; SEQ ID NO 34
; LENGTH: 2321
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020192678A1 1382924.50
US-10-071-766-34

Query Match          0.6%; Score 18; DB 9; Length 2321;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1063 GAAAAGATGCTCAAAA 1080
Db 2294 GAAAAGATGCTCAAAA 2311
```

## RESULT 12

US-09-815-242-7983  
; Sequence 7983, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Karl L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; FILE REFERENCE: ELITRA.011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7983  
; LENGTH: 4446  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(4446)  
US-09-815-242-7983

Query Match 0.6%; Score 18; DB 10; Length 4446;  
Best Local Similarity 100.0%; Pred. No. 30;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1819 ATGACCGCGCGCAAGCTG 1836  
|||||  
Db 3904 ATGACCGCGCGCAAGCTG 3921

## RESULT 13

US-09-070-927A-238/c  
; Sequence 238, Application US/09070927A  
; Patent No. US20020120116A1  
; GENERAL INFORMATION:  
; APPLICANT: Charles A. Kunsch  
; Patrick J. Dillon  
; Steven Barash  
; TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides  
; NUMBER OF SEQUENCES: 982  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/070,927A  
; FILING DATE: 04-May-2000  
; CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/046,655  
; FILING DATE: 1997-05-16  
; APPLICATION NUMBER: 60/044,031  
; FILING DATE: 1997-05-06  
; APPLICATION NUMBER: 60/066,009  
; FILING DATE: 1997-11-14  
ATTORNEY/AGENT INFORMATION:  
; NAME: Kenley K. Hoover  
; REGISTRATION NUMBER: 40,302  
; REFERENCE/DOCKET NUMBER: PB369  
TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 238:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 29729 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; SEQUENCE DESCRIPTION: SEQ ID NO: 238:  
US-09-070-927A-238

Query Match 0.6%; Score 18; DB 10; Length 29729;  
Best Local Similarity 100.0%; Pred. No. 32;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1064 AAAAAGATGCTCAAAAAG 1081  
|||||  
Db 9605 AAAAAGATGCTCAAAAAG 9588

## RESULT 14

US-09-789-836-22  
; Sequence 22, Application US/09789836  
; Patent No. US20020082204A1  
; GENERAL INFORMATION:  
; APPLICANT: BRIGHAM, KENNETH L.  
; APPLICANT: STECENKO, ARLENE A.  
; APPLICANT: SEALY, LINDA  
; TITLE OF INVENTION: TREATMENT OF INFLAMMATION WITH P20  
; FILE REFERENCE: N-6977  
; CURRENT APPLICATION NUMBER: US/09/789,836  
; CURRENT FILING DATE: 2001-02-20  
; PRIOR APPLICATION NUMBER: 60/183,584  
; PRIOR FILING DATE: 2000-02-18  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 22  
; LENGTH: 102  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:  
; OTHER INFORMATION: Oligonucleotide  
US-09-789-836-22

Query Match 0.5%; Score 17; DB 10; Length 102;  
Best Local Similarity 100.0%; Pred. No. 87;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 87 AGCGCGCGCGCGGCA 103  
|||||  
Db 23 AGCGCGCGCGCGGCA 39

## RESULT 15

US-09-789-836-21/c

; Sequence 21, Application US/09789836  
; Patent No. US20020082204A1  
; GENERAL INFORMATION:  
; APPLICANT: BRIGHAM, KENNETH L.  
; APPLICANT: STECENKO, ARLENE A.  
; APPLICANT: SEALY, LINDA

; TITLE OF INVENTION: TREATMENT OF INFLAMMATION WITH P20

; FILE REFERENCE: N-6977

; CURRENT APPLICATION NUMBER: US/09/789,836

; CURRENT FILING DATE: 2001-02-20

; PRIOR APPLICATION NUMBER: 60/183,584

; PRIOR FILING DATE: 2000-02-18

; NUMBER OF SEQ ID NOS: 33

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 21

; LENGTH: 105

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence:

; OTHER INFORMATION: Oligonucleotide

US-09-789-836-21

Query Match 0.5%; Score 17; DB 10; Length 105;

Best Local Similarity 100.0%; Pred. No. 87;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 87 AGCGCGCGCGCGGCA 103

|||||

Db 83 AGCGCGCGCGCGGCA 67

RESULT 16

US-09-789-836-33/c

; Sequence 33, Application US/09789836

; Patent No. US20020082204A1

; GENERAL INFORMATION:

; APPLICANT: BRIGHAM, KENNETH L.

; APPLICANT: STECENKO, ARLENE A.

; APPLICANT: SEALY, LINDA

; TITLE OF INVENTION: TREATMENT OF INFLAMMATION WITH P20

; FILE REFERENCE: N-6977

; CURRENT APPLICATION NUMBER: US/09/789,836

; CURRENT FILING DATE: 2001-02-20

; PRIOR APPLICATION NUMBER: 60/183,584

; PRIOR FILING DATE: 2000-02-18

; NUMBER OF SEQ ID NOS: 33

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 33

; LENGTH: 106

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence:

; OTHER INFORMATION: Oligonucleotide

US-09-789-836-33

Query Match 0.5%; Score 17; DB 10; Length 106;

Best Local Similarity 100.0%; Pred. No. 87;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 87 AGCGCGCGCGCGGCA 103

|||||

Db 84 AGCGCGCGCGCGGCA 68

RESULT 17

US-09-783-590-5143

; Sequence 5143, Application US/09783590

; Patent No. US20020110850A1

; GENERAL INFORMATION:

; APPLICANT: DILLON, Patrick J.

; APPLICANT: Haseltine, William A.

; APPLICANT: Li, Haodong  
; APPLICANT: Rosen, Craig A.  
; APPLICANT: Ruben, Steven M.  
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2  
; FILE REFERENCE: PO-16.2C1

; CURRENT APPLICATION NUMBER: US/09/783,590

; CURRENT FILING DATE: 2000-02-15

; PRIOR APPLICATION NUMBER: 08/420,856

; PRIOR FILING DATE: 1995-04-12

; PRIOR APPLICATION NUMBER: 08/346,731

; PRIOR FILING DATE: 1994-11-21

; NUMBER OF SEQ ID NOS: 12485

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 5143

; LENGTH: 122

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (15)

; OTHER INFORMATION: n equals a,t,g, or c

; NAME/KEY: misc feature

; LOCATION: (60)

; OTHER INFORMATION: n equals a,t,g, or c

; NAME/KEY: misc feature

; LOCATION: (109)

; OTHER INFORMATION: n equals a,t,g, or c

; NAME/KEY: misc feature

; LOCATION: (120)

; OTHER INFORMATION: n equals a,t,g, or c

US-09-783-590-5143

Query Match

Best Local Similarity 0.5%; Score 17; DB 10; Length 122;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GCGCGCGCGCGGCAC 104

|||||

Db 35 GCGCGCGCGCGGCAC 51

RESULT 18

US-09-974-300-4131/c

; Sequence 4131, Application US/09974300

; Patent No. US20020146721A1

; GENERAL INFORMATION:

; APPLICANT: Berk, Randy M.

; APPLICANT: Clausen, Ib Groth

; TITLE OF INVENTION: Methods For Monitoring Multiple Gene

; FILE REFERENCE: Expression

; FILE REFERENCE: 10085.500-US

; CURRENT APPLICATION NUMBER: US/09/974,300

; CURRENT FILING DATE: 2001-10-05

; PRIOR APPLICATION NUMBER: 09/680,598

; PRIOR FILING DATE: 2000-10-06

; PRIOR APPLICATION NUMBER: 60/279,526

; PRIOR FILING DATE: 2001-03-27

; NUMBER OF SEQ ID NOS: 8481

; SOFTWARE: FastSeq for Windows Version 4.0.

; SEQ ID NO 4131

; LENGTH: 126

; TYPE: DNA

; ORGANISM: Bacillus licheniformis

US-09-974-300-4131

Query Match

Best Local Similarity 0.5%; Score 17; DB 10; Length 126;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1510 AAAAAGCGCGGCGCA 1526

|||||

Db 94 AAAAAGCGCGGCGCA 78



```
RESULT 19
US-09-878-574-8654
; Sequence 8654, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 8654
; LENGTH: 224
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: 701101486H1
US-09-878-574-8654

Query Match      0.5%; Score 17; DB 10; Length 224;
Best Local Similarity 100.0%; Pred. No. 90;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1353 CGTGGACAGCAAGTTTCG 1369
      |||
Db 172 CGTGGACAGCAAGTTTCG 188

RESULT 20
US-09-880-107-3632
; Sequence 3632, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3632
; LENGTH: 335
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 W87606
US-09-880-107-3632

Query Match      0.5%; Score 17; DB 10; Length 335;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1258 ACGGCGGCTCTGCTGCT 1274
      |||
Db 120 ACGGCGGCTCTGCTGCT 136

RESULT 21
US-09-974-300-7591/c
; Sequence 7591, Application US/09974300
; Patent No. US20020146721A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; TITLE OF INVENTION: Expression
; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7591
; LENGTH: 336
; TYPE: DNA
; ORGANISM: Bacillus clausii
US-09-974-300-7591

Query Match      0.5%; Score 17; DB 10; Length 336;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2768 ACGTCCGTTTGC CGCA 2784
      |||
Db 115 ACGTCCGTTTGC CGCA 99

RESULT 22
US-09-983-965-4734/C
; Sequence 4734, Application US/09983965
; Patent No. US20020137160A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 37-21(10297)C
; CURRENT APPLICATION NUMBER: US/09/983,965
; CURRENT FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: US 09/465,231
; PRIOR FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: US 60/113,678
; PRIOR FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 5912
; SEQ ID NO 4734
; LENGTH: 381
; TYPE: DNA
; ORGANISM: Bos taurus
; FEATURE:
; OTHER INFORMATION: Clone ID: 17-LIB34-006-Q1-E1-E9
US-09-983-965-4734

Query Match      0.5%; Score 17; DB 10; Length 381;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 524 AAGACGGAGCGGTAAA 540
      |||
Db 70 AAGACGGAGCGGTAAA 54

RESULT 23
US-10-040-739-1122/c
; Sequence 1122, Application US/10040739
; Patent No. US20020173635A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John
; APPLICANT: LaVallie, Edward
```

```

; FEATURE:
; OTHER INFORMATION: Clone ID: 33-LIB34-008-Q1-E1-A2
US-09-983-965-5116

Query Match          0.5%; Score 17; DB 10; Length 415;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 17; Conservative 0; Mismatches 0; Indels

QY 156 AGCAACAACAGCGAAAT 172
      |||||
Db 398 AGCAACAACAGCGAAAT 414

RESULT 25
US-09-854-133-339/c
; Sequence 339, Application US/09854133
; Publication No. US20020183499A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Henderson, Robert A.
; APPLICANT: Benson, Darin R.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C10
; CURRENT APPLICATION NUMBER: US/09/854.133
; CURRENT FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 735
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 339
; LENGTH: 440
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(440)
; OTHER INFORMATION: n = A,T,C or G
US-09-854-133-339

Query Match          0.5%; Score 17; DB 9; Length 440;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 17; Conservative 0; Mismatches 0; Indels

QY 87 AGCGCGCGCGCGCGCA 103
      |||||
Db 59 AGCGCGCGCGCGCGCA 43

RESULT 26
US-09-738-973-339/c
; Sequence 339, Application US/09738973
; Patent No. US20020110563A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Fling, Steven P.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Algate, Paul A.
; APPLICANT: Secrist, Heather
; APPLICANT: Indirias, Carol Yoseph
; APPLICANT: Benson, Darin R.
; APPLICANT: Elliott, Mark
; APPLICANT: Mannion, Jane
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C9
; CURRENT APPLICATION NUMBER: US/09/738.973
; CURRENT FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 587
; SOFTWARE: FastSeq for Windows Version 3.0

```

; SEQ ID NO 339  
; LENGTH: 440  
; TYPE: DNA  
; ORGANISM: Homo sapien  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(440)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-738-973-339

Query Match 0.5%; Score 17; DB 10; Length 440;  
Best Local Similarity 100.0%; Pred. No. 92;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 87 AGCGCGCGCGCGGCA 103  
|||||

Db 59 AGCGCGCGCGCGGCA 43

RESULT 27

; Sequence 7839, Application US/09815242  
; Patent No. US20020061569A1

; GENERAL INFORMATION:

; APPLICANT: Haseibeck, Robert  
; APPLICANT: Ohlsen, Karl L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard

; TITLE OF INVENTION: Identification of Essential Genes in  
; FILE REFERENCE: ELITRA.011A

; CURRENT APPLICATION NUMBER: US/09/815,242  
; CURRENT FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 7839

; LENGTH: 441

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)...(441)

US-09-815-242-7839

Query Match 0.5%; Score 17; DB 10; Length 441;  
Best Local Similarity 100.0%; Pred. No. 92;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1524 CCAACTGCAACTGCACG 1540  
|||||

Db 315 CCAACTGCAACTGCACG 331

RESULT 28

US-09-864-761-204/c

; Sequence 204, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO  
; FILE REFERENCE: Aeomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1

; SEQ ID NO 204

; LENGTH: 480

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: MAP TO AL049780.2

; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5

; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.2

; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5

; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.7

; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2

; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.5

; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.4

; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7

; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 3.9

; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.6

US-09-864-761-204

Query Match 0.5%; Score 17; DB 10; Length 480;  
Best Local Similarity 100.0%; Pred. No. 92;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2595 CATTAGTCTGTTGCAG 2611

```
Db 278 CATTAGTCTGTTGCAG 262
|||||
; TITLE OF INVENTION: Encoding No. US20020115078A1el Helicobacter Polypeptides in t
; TITLE OF INVENTION: Genome
; FILE REFERENCE: 06132/041002
; CURRENT APPLICATION NUMBER: US/09/881,752A
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 08/833,457
; PRIOR FILING DATE: 1997-04-01
; NUMBER OF SEQ ID NOS: 370
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 225
; LENGTH: 490
; TYPE: DNA
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (51)...(437)
US-09-881-752A-225

Query Match 0.5%; Score 17; DB 10; Length 490;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 991 ATGCTTTTCATTTTTC 1007
Db 61 ATGCTTTTCATTTTTC 45
|||||

RESULT 31
US-09-864-761-8567/c
; Sequence 8567, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/006666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21

US-10-040-739-17/c
; Sequence 17, Application US/10040739
; Patent No. US20020173635A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John
; APPLICANT: Lavalie, Edward
; APPLICANT: Racie, Lisa
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Spaulding, Vikki
; TITLE OF INVENTION: SECRETED, EXPRESSED SEQUENCE TAGS
; NUMBER OF SEQUENCES: 1519
; CORRESPONDENCE ADDRESS:
; ADDRESS: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/10/040,739
; FILING DATE: 07-Jan-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/036,520
; FILING DATE: 03-JUN-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Brown, Scott A.
; REGISTRATION NUMBER: 32,724
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8224
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 484 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-10-040-739-17

Query Match 0.5%; Score 17; DB 9; Length 484;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 741 CGAAATAATGCTGCAG 757
Db 299 CGAAATAATGCTGCAG 283
|||||

RESULT 30
US-09-881-752A-225/c
; Sequence 225, Application US/09881752A
; Patent No. US20020115078A1
; GENERAL INFORMATION:
; APPLICANT: Kleanthous, Harold
; APPLICANT: Al-Garawi, Amal
; APPLICANT: Miller, Charles
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Oomen, Raymond P.
; TITLE OF INVENTION: Identification of Polynucleotides
```



US-09-738-626-2896  
; Sequence 2896, Application US/09738626  
; Publication No. US20020197605A1  
; GENERAL INFORMATION:  
; APPLICANT: NAKAGAWA, SATOSHI  
; APPLICANT: MIZOGUCHI, HIROSHI  
; APPLICANT: ANDO, SEIKO  
; APPLICANT: HAYASHI, MIKIRO  
; APPLICANT: OCHIAI, KEIKO  
; APPLICANT: YOKOI, HARUHIKO  
; APPLICANT: TATEISHI, NAOKO  
; APPLICANT: SENOH, AKIHIRO  
; APPLICANT: IKEDA, MASATO  
; APPLICANT: OZAKI, AKIO  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-125  
; CURRENT APPLICATION NUMBER: US/09/738,626  
; CURRENT FILING DATE: 2000-12-18  
; PRIOR APPLICATION NUMBER: JP 99/377484  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: JP 00/159162  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: JP 00/280988  
; PRIOR FILING DATE: 2000-08-03  
; NUMBER OF SEQ ID NOS: 7059  
; SOFTWARE: PatentIn ver. 3.0  
; SEQ ID NO 2896  
; LENGTH: 747  
; TYPE: DNA  
; ORGANISM: Corynebacterium glutamicum  
US-09-738-626-2896

Query Match 0.5%; Score 17; DB 9; Length 747;  
Best Local Similarity 100.0%; Pred. No. 94;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 2401 GGTCTGCGCGTCATCGC 2417  
Db 37 GGTCTGCGCGTCATCGC 53

RESULT 35  
US-09-764-864-326/c  
; Sequence 326, Application US/09764864  
; Patent No. US20020132753A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT223  
; CURRENT APPLICATION NUMBER: US/09/764,864  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 1792  
; SOFTWARE: PatentIn ver. 2.0  
; SEQ ID NO 326  
; LENGTH: 872  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-764-864-326

Query Match 0.5%; Score 17; DB 10; Length 872;  
Best Local Similarity 100.0%; Pred. No. 94;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 88 GCGCGCGCGCGCGCAC 104  
Db 105 GCGCGCGCGCGCGCAC 89

RESULT 36  
US-09-938-842A-2261  
; Sequence 2261, Application US/09938842A  
; Patent No. US20020160378A1

; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; FILE REFERENCE: SAME, AND METHODS OF USE  
; FILE REFERENCE: SCRIPL300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; CURRENT FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 2261  
; LENGTH: 915  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-2261

Query Match 0.5%; Score 17; DB 9; Length 915;  
Best Local Similarity 100.0%; Pred. No. 95;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 2456 GTGTTGAAGGCAAAATG 2472  
Db 317 GTGTTGAAGGCAAAATG 333

RESULT 37  
US-09-974-300-5167  
; Sequence 5167, Application US/09974300  
; Patent No. US20020146721A1  
; GENERAL INFORMATION:  
; APPLICANT: Berka, Randy M.  
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene  
; TITLE OF INVENTION: Expression  
; FILE REFERENCE: 10085.500-US  
; CURRENT APPLICATION NUMBER: US/09/974,300  
; CURRENT FILING DATE: 2001-10-05  
; PRIOR APPLICATION NUMBER: 09/680,598  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/279,526  
; PRIOR FILING DATE: 2001-03-27  
; NUMBER OF SEQ ID NOS: 8481  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5167  
; LENGTH: 960  
; TYPE: DNA  
; ORGANISM: Bacillus clausii  
US-09-974-300-5167

Query Match 0.5%; Score 17; DB 10; Length 960;  
Best Local Similarity 100.0%; Pred. No. 95;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 2944 AAAGCGCTCTGTTTGC 2960  
Db 333 AAAGCGCTCTGTTTGC 349

RESULT 38  
US-09-789-836-2/c  
; Sequence 2, Application US/09789836  
; Patent No. US20020082204A1  
; GENERAL INFORMATION:  
; APPLICANT: BRIGHAM, KENNETH L.  
; APPLICANT: STECENKO, ARLENE A.  
; APPLICANT: SEALY, LINDA

1  
; TITLE OF INVENTION: TREATMENT OF INFLAMMATION WITH P20  
; FILE REFERENCE: N-6977  
; CURRENT APPLICATION NUMBER: US/09/789,836  
; CURRENT FILING DATE: 2001-02-20  
; PRIOR APPLICATION NUMBER: 60/183,584  
; PRIOR FILING DATE: 2000-02-18  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 1038  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-789-836-2

Query Match 0.5%; Score 17; DB 10; Length 1038;  
Best Local Similarity 100.0%; Pred. No. 95;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 87 AGCGCGCGCGCGGCA 103  
|||||  
Db 57 AGCGCGCGCGCGGCA 41

RESULT 39  
US-09-822-830A-366/c  
; Sequence 366, Application US/09822830A  
; Patent No. US20020142952A1  
; GENERAL INFORMATION:  
; APPLICANT: Genetics Institute, Inc.  
; APPLICANT: Wong, Gordon G.  
; APPLICANT: Clark, Hilary  
; APPLICANT: Fechtel, Kim  
; APPLICANT: Agostino, Michael J.  
; APPLICANT: Howes, Steven H.  
; APPLICANT: Resnick, Richard J.  
; APPLICANT: Gulukota, Kamalakara  
; APPLICANT: Graham, James R.  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS  
; FILE REFERENCE: GIN 6402  
; CURRENT APPLICATION NUMBER: US/09/822,830A  
; CURRENT FILING DATE: 2001-03-29  
; PRIOR APPLICATION NUMBER: 60/195,604  
; PRIOR FILING DATE: 2000-04-06  
; NUMBER OF SEQ ID NOS: 631  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 366  
; LENGTH: 1408  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: 1254..1257  
; OTHER INFORMATION: n=a,c,g, or t  
US-09-822-830A-366

Query Match 0.5%; Score 17; DB 10; Length 1408;  
Best Local Similarity 100.0%; Pred. No. 96;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2748 GGGCGCACTGGCGGTG 2764  
|||||  
Db 922 GGGCGCACTGGCGGTG 906

RESULT 40  
US-10-108-915-33/c  
; Sequence 33, Application US/10108915  
; Patent No. US2002017204A1  
; GENERAL INFORMATION:  
; APPLICANT: Cahoon, Rebecca E.  
; APPLICANT: Shen, Jennie  
; APPLICANT: Williams, Mark  
; TITLE OF INVENTION: Geranylgeranyl Pyrophosphate Synthases

; FILE REFERENCE: BB1286 US NA  
; CURRENT APPLICATION NUMBER: US/10/108,915  
; CURRENT FILING DATE: 2002-03-28  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/452,238  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-01  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/110,592  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-December-02  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 33  
; LENGTH: 1428  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
US-10-108-915-33

Query Match 0.5%; Score 17; DB 9; Length 1428;  
Best Local Similarity 100.0%; Pred. No. 96;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 87 AGCGCGCGCGCGGCA 103  
|||||  
Db 169 AGCGCGCGCGCGGCA 153

Search completed: January 28, 2003, 00:31:35  
Job time : 250 secs





GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 27, 2003, 19:09:24 ; Search time 433 Seconds

(without alignments)

16663.747 Million cell updates/sec

Title: US-09-830-433A-7

Perfect score: 3204

Sequence: 1 atgcgaacgaccccaacctt.....gcgtaggctaccggtctga 3204

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 1125999159 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : N\_Geneseq\_101002.\*

- 1: /SID52/qcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*
- 2: /SID52/qcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*
- 3: /SID52/qcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*
- 4: /SID52/qcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*
- 5: /SID52/qcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*
- 6: /SID52/qcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*
- 7: /SID52/qcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*
- 8: /SID52/qcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*
- 9: /SID52/qcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*
- 10: /SID52/qcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*
- 11: /SID52/qcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*
- 12: /SID52/qcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*
- 13: /SID52/qcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*
- 14: /SID52/qcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*
- 15: /SID52/qcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*
- 16: /SID52/qcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*
- 17: /SID52/qcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*
- 18: /SID52/qcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*
- 19: /SID52/qcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*
- 20: /SID52/qcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*
- 21: /SID52/qcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*
- 22: /SID52/qcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*
- 23: /SID52/qcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*
- 24: /SID52/qcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	3204	100.0	3204	21	AAA15299 DNA encoding a pol
2	990	30.9	1365	21	AAZ53301 Neisseria meningit
3	834	26.0	1365	21	AAZ53300 Neisseria meningit
4	834	26.0	4218	22	AAZ53905 Neisseria meningit
5	834	26.0	4218	22	AAZ17058 N. meningitidis st
6	834	26.0	13423	21	AAA81527 N. meningitidis pa
7	834	26.0	349980	21	AAZ1612 Neisseria meningit
8	831	25.9	3939	22	AAZ3875 Neisseria meningit
9	831	25.9	3939	22	AAZ3880 Neisseria meningit

10	831	25.9	3939	22	AAZ17040 N. meningitidis st
11	831	25.9	3939	22	AAZ17045 N. meningitidis st
12	831	25.9	4170	22	AAZ43902 Neisseria meningit
13	831	25.9	4170	22	AAZ17055 N. meningitidis st
14	831	25.9	4179	22	AAZ43877 Neisseria meningit
15	831	25.9	4179	22	AAZ17042 N. meningitidis st
16	831	25.9	4335	22	AAZ43899 Neisseria meningit
17	831	25.9	4335	22	AAZ17052 N. meningitidis st
18	831	25.9	4344	22	AAZ43876 Neisseria meningit
19	831	25.9	4425	22	AAZ43874 Neisseria meningit
20	831	25.9	4425	22	AAZ17039 N. meningitidis st
21	398	12.4	3798	22	AAZ17041 N. meningitidis st
22	189	5.9	526	21	AAZ53298 Neisseria meningit
23	177	5.5	286	19	AAV03603 Neisseria meningit
24	177	5.5	286	19	AAV03603 Genomic fragment o
25	160	5.0	528	21	AAZ53297 Neisseria meningit
26	149	4.7	537	21	AAZ53296 Neisseria gonorrhe
27	131	4.1	1365	21	AAZ53299 Neisseria gonorrhe
c	29	1.8	476	21	AAA81823 N. meningitidis pa
29	41	1.3	577	21	AAZ82003 N. meningitidis pa
30	35	1.1	50	22	AAZ44248 Neisseria meningit
31	25	0.8	2280	23	ABL05527 Drosophila melanog
c	32	0.8	3302	23	ABL28074 Drosophila melanog
c	33	0.8	13327	23	ABL05526 Drosophila melanog
c	34	0.7	30	21	AAA15341 PCR primer for a N
35	22	0.7	2382	23	ABL27930 Drosophila melanog
36	21	0.7	34	21	AAA15340 PCR primer for a N
37	20	0.6	34	21	AAZ55456 Neisseria species
c	38	0.6	735	18	AAZ45009 Serratia marcescen
c	39	0.6	735	19	AAV43062 DNA encoding an ac
c	40	0.6	1105	24	ABQ32992 Oligonucleotide fo
41	20	0.6	1105	24	ABQ32993 Oligonucleotide fo
c	42	0.6	5019	22	AAF82427 Moraxella catarrha
43	20	0.6	15004	23	ABL27628 Drosophila melanog
c	44	0.6	89047	22	AAF28547 Genomic fragment #
45	19	0.6	29	22	AAZ44242 Neisseria meningit
46	32	0.6	284	24	AAZ44246 Neisseria meningit
47	19	0.6	284	24	ABV14705 Human ovarian canc
c	48	0.6	455	23	ABV14705 Human prostate exp
c	49	0.6	570	22	AAH12092 Human cDNA clone (
50	19	0.6	807	22	AAH52244 Human AFP protein
51	19	0.6	1267	21	AAC98115 Human colon cancer
52	19	0.6	1286	22	AAI93900 Human stomach canc
53	19	0.6	1286	22	AAH16346 Human cDNA sequenc
54	19	0.6	1383	22	AAH52253 Human AFP protein
55	19	0.6	1451	22	AAH3753 Human cDNA sequenc
c	56	0.6	1551	24	ABN95857 Gene #2355 used to
57	19	0.6	1596	22	AAH52254 Human AFP protein
58	19	0.6	1611	22	AAK77974 Human digestive sy
59	19	0.6	1987	21	AAC77912 Human cancer assoc
60	19	0.6	2000	13	AAO25783 CTP synthetase gene
61	19	0.6	3001	12	AAQ14668 CTP synthetase gen
62	19	0.6	3017	24	ABL55297 Human mitochondria
63	19	0.6	4424	22	AAK51737 Human polynucleoti
64	19	0.6	4504	22	AAK52721 Human polynucleoti
c	65	0.6	11340	23	ABL16852 Drosophila melanog
66	19	0.6	17687	22	AAZ42069 Genomic sequence #
67	19	0.6	19815	22	AAZ42064 Genomic sequence #
68	19	0.6	62909	22	AAF28545 Genomic fragment #
c	69	0.6	28	22	AAZ44225 Neisseria meningit
c	70	0.6	28	22	AAZ17068 N. meningitidis fu
c	71	0.6	30	22	AAZ15111 Extended_TOGA prim
c	72	0.6	60	16	AAT00248 Thrombin 60N DNA 1
c	73	0.6	60	22	AAZ70800 Thrombin high affi
74	18	0.6	111	22	AAZ6369 P. putida oxygenas
75	18	0.6	213	24	ABK77439 Bacillus clausii g
c	76	0.6	267	20	AAV90490 EST clone DN714.
c	77	0.6	267	20	AAV89054 EST clone CB123.
78	18	0.6	288	21	AAC76303 Human OREF1858
c	79	0.6	378	15	AAT24092 Human gene signatu
c	80	0.6	504	23	ABV48468 Human prostate exp
81	18	0.6	507	24	ABQ30826 Oligonucleotide fo
c	82	0.6	507	24	ABQ30827 Oligonucleotide fo

c 83	18	0.6	592	22	AAH09157	Human cDNA clone (	c 156	17	0.5	54	20	AAZ00332	Mutated hTR promot
c 84	18	0.6	593	24	ABQ51880	Oligonucleotide fo	c 157	17	0.5	77	17	AAAT09808	Tomato genomic DNA
c 85	18	0.6	593	24	ABQ51881	Oligonucleotide fo	c 158	17	0.5	102	22	AAAD19388	Synthetic mutant
c 86	18	0.6	633	23	AAH67168	DNA encoding novel	c 159	17	0.5	104	23	ABK51408	Human CCAAT enhanc
c 87	18	0.6	653	22	AAH34289	Human colon cancer	c 160	17	0.5	106	22	AAAD19387	Synthetic mutant t
c 88	18	0.6	720	24	ABQ33460	Oligonucleotide fo	c 161	17	0.5	106	22	AAAD19399	Wild type C/EBPbet
c 89	18	0.6	720	24	ABQ33461	Oligonucleotide fo	c 162	17	0.5	107	22	AAZ07325	Mutated hTR promot
c 90	18	0.6	797	24	ABQ46156	Oligonucleotide fo	c 163	17	0.5	116	22	AAAS39798	Genomic sequence #
c 91	18	0.6	797	24	ABQ46157	Oligonucleotide fo	c 164	17	0.5	116	22	AAAK90151	Human digestive sy
c 92	18	0.6	885	24	AAAD24396	Human RNA metaboli	c 165	17	0.5	117	18	AAAT93339	EST locus y82el2.
c 93	18	0.6	888	22	AAH24356	Human ribosome sub	c 166	17	0.5	126	24	ABK76840	Bacillus lichenifo
c 94	18	0.6	962	21	AAAC93388	Human secreted pro	c 167	17	0.5	134	21	AAAC09407	Human telomerase R
c 95	18	0.6	994	20	AAV84522	Human secreted pro	c 168	17	0.5	176	20	AAZ07323	Human telomerase R
c 96	18	0.6	994	22	ABAB3305	Human secreted pro	c 169	17	0.5	267	24	AAZ07251	Mouse digital sequ
c 97	18	0.6	1044	24	ABK75024	Bacillus lichenifo	c 170	17	0.5	267	24	AAAD37707	Mouse digestive sequ
c 98	18	0.6	1104	23	ABL24461	Drosophila melanog	c 171	17	0.5	289	24	ABI99576	Mouse ischaemic co
c 99	18	0.6	1136	22	AAI59751	Human polynucleoti	c 172	17	0.5	315	24	ABN79072	Human ORF4019 cDNA
c 100	18	0.6	1139	21	AAAC76370	Human polynucleoti	c 173	17	0.5	317	22	AAH69137	Human cervical can
c 101	18	0.6	1169	22	AAI57965	Human polynucleoti	c 174	17	0.5	320	22	AAI89249	Human polynucleoti
c 102	18	0.6	1177	22	AAAF32742	Human secreted pro	c 175	17	0.5	322	18	AAAT93340	EST locus yx4f06
c 103	18	0.6	1183	24	AAAL40191	Isoprenoid related	c 176	17	0.5	335	24	ABN97137	Gene #3635 used to
c 104	18	0.6	1335	23	ABL22177	Drosophila melanog	c 177	17	0.5	336	24	ABK80300	Bacillus clausii g
c 105	18	0.6	1375	21	AAZ33365	Human secreted pro	c 178	17	0.5	347	22	AAK84364	Human immune/haema
c 106	18	0.6	1431	24	ABQ42300	Oligonucleotide fo	c 179	17	0.5	352	22	AAAF65999	Novel human polynu
c 107	18	0.6	1431	24	ABQ42301	Oligonucleotide fo	c 180	17	0.5	356	21	AAAC10429	Human secreted pro
c 108	18	0.6	1434	23	AAAS80223	DNA encoding novel	c 181	17	0.5	362	21	AAAC57588	Arachidonic acid m
c 109	18	0.6	1501	24	ABQ43092	Oligonucleotide fo	c 182	17	0.5	362	21	AAAC57589	Arachidonic acid m
c 110	18	0.6	1501	24	ABQ43093	Oligonucleotide fo	c 183	17	0.5	362	21	AAAC57590	Arachidonic acid m
c 111	18	0.6	1531	23	AAAS84360	DNA encoding novel	c 184	17	0.5	362	21	AAAC57591	Arachidonic acid m
c 112	18	0.6	1614	22	AAH17079	Human cDNA sequenc	c 185	17	0.5	363	22	AAI82582	Human polynucleoti
c 113	18	0.6	1689	22	AAAF80591	Receptor #79 parti	c 186	17	0.5	376	22	AAI83211	Human polynucleoti
c 114	18	0.6	1721	21	AAAC79728	Human secreted pro	c 187	17	0.5	384	21	AAAC74894	Human OREF ORF449
c 115	18	0.6	1903	19	AAAX14490	H. pylori GHPO 113	c 188	17	0.5	388	20	AAV88644	EST clone GV307.
c 116	18	0.6	2363	20	AAAT78215	Melon MEERS cDNA.	c 189	17	0.5	396	21	AAAC04103	Human secreted pro
c 117	18	0.6	2464	22	AAAS52213	cDNA encoding nove	c 190	17	0.5	400	18	AAV77284	Staphylococcus aur
c 118	18	0.6	2536	24	ABK35738	cDNA sequence #129	c 191	17	0.5	404	23	ABV03564	Human prostate exp
c 119	18	0.6	2682	22	AAH14265	Human cDNA sequenc	c 192	17	0.5	419	21	AAAC09275	Human secreted pro
c 120	18	0.6	2726	23	ABL09441	Drosophila melanog	c 193	17	0.5	420	23	ABV12733	Human prostate exp
c 121	18	0.6	2902	22	AAAC90433	Murine lymph node	c 194	17	0.5	429	21	ABQ63154	Mycobacterium tube
c 122	18	0.6	2905	23	AAAT78259	DNA encoding novel	c 195	17	0.5	440	21	AAAC79297	Human lung tumour-
c 123	18	0.6	2943	23	AAAS86435	DNA encoding novel	c 196	17	0.5	440	21	AAAC79297	Human lung tumour-
c 124	18	0.6	3030	23	ABL17899	Drosophila melanog	c 197	17	0.5	441	23	AAAS4202	Pseudomonas aerugi
c 125	18	0.6	3281	23	ABL09450	Drosophila melanog	c 198	17	0.5	443	23	ABV18440	Human prostate exp
c 126	18	0.6	3766	23	ABL22176	Drosophila melanog	c 199	17	0.5	447	23	AAAS81595	DNA encoding novel
c 127	18	0.6	3971	23	ABL17923	Drosophila melanog	c 200	17	0.5	463	23	ABV33862	Human prostate exp
c 128	18	0.6	4446	23	AAAS4346	Pseudomonas aerugi	c 201	17	0.5	463	23	ABV42747	Human prostate exp
c 129	18	0.6	5080	23	ABL09440	Drosophila melanog	c 202	17	0.5	463	23	ABV42747	Human prostate exp
c 130	18	0.6	5893	24	ABL32859	Human immune syste	c 203	17	0.5	480	22	ABA51921	Human foetal liver
c 131	18	0.6	5952	24	AAAS61403	Human gene regulat	c 204	17	0.5	480	22	ABA21738	Probe #204 for gen
c 132	18	0.6	6988	23	ABL29124	Drosophila melanog	c 205	17	0.5	480	22	AAK00215	Human brain expres
c 133	18	0.6	6999	23	AAAS85200	DNA encoding novel	c 206	17	0.5	480	22	AAK25655	Human bone marrow
c 134	18	0.6	7167	24	ABL32400	Human immune syste	c 207	17	0.5	480	22	AAI10283	Probe #216 for gen
c 135	18	0.6	7800	23	ABL24460	Drosophila melanog	c 208	17	0.5	480	22	AAI131531	Probe #217 used to
c 136	18	0.6	7875	21	AAA92470	Shewanella sp. SCR	c 209	17	0.5	480	22	AAH46893	cDNA encoding huma
c 137	18	0.6	9299	23	ABL33725	Human immune syste	c 210	17	0.5	480	24	ABS00227	Human genome-deriv
c 138	18	0.6	10337	23	ABL25704	Drosophila melanog	c 211	17	0.5	484	20	AAV87539	EST clone DD426.
c 139	18	0.6	12986	22	AAAS46553	Tumour suppressor	c 212	17	0.5	487	22	AAK92117	Human cDNA 5'-end
c 140	18	0.6	14542	24	ABL70191	Chemically treated	c 213	17	0.5	487	22	AAK93380	Human cDNA clone r
c 141	18	0.6	14542	24	AAAS61147	Human gene regulat	c 214	17	0.5	490	19	AAAI4032	H. pylori GHPO 893
c 142	18	0.6	14542	24	ABK31234	Signal transductio	c 215	17	0.5	490	22	AAAL04542	Human reproductive
c 143	18	0.6	16560	23	ABL17922	Drosophila melanog	c 216	17	0.5	490	23	ABL97465	Human testicular a
c 144	18	0.6	16759	24	ABL36297	Human lysosomal ac	c 217	17	0.5	494	23	ABV48225	Human prostate exp
c 145	18	0.6	17893	24	ABL33365	Human immune syste	c 218	17	0.5	500	18	AAAT43451	ATM gene exon 11.
c 146	18	0.6	19165	23	ABL17898	Drosophila melanog	c 219	17	0.5	500	22	AAAF57603	ATM genomic fragme
c 147	18	0.6	23821	21	AAA92469	Shewanella sp. SCR	c 220	17	0.5	518	22	AAK61544	Human immune/haema
c 148	18	0.6	29729	20	AAAX13175	Enterococcus faeca	c 221	17	0.5	521	21	AAA98391	Red buffalo PTFN/M
c 149	18	0.6	50925	21	AAA81487	N. meningitidis pa	c 222	17	0.5	524	24	ABQ14858	Human testicular a
c 150	18	0.6	349980	21	AAF21610	Weissaria meningit	c 223	17	0.5	524	24	ABQ14859	Oligonucleotide fo
c 151	18	0.6	1437668	21	AAAB1480	N. meningitidis B	c 224	17	0.5	526	23	ABV58421	Human prostate exp
c 152	17	0.5	20	24	AAAD37756	Real-time validati	c 225	17	0.5	532	22	ABA62815	Human foetal liver
c 153	17	0.5	51	22	AAAL28807	Human SNP oligonuc	c 226	17	0.5	532	22	ABA30101	Probe #8567 for ge
c 154	17	0.5	54	20	AAZ07332	Mutated hTR promot	c 227	17	0.5	532	22	AAK11214	Human brain expres
c 155	17	0.5	54	20	AAZ00333	Mutated hTR promot	c 228	17	0.5	532	22	AAK37015	Human bone marrow

c 229	17	0.5	532	22	AA117860	Probe #7793 for ge	c 302	17	0.5	890	24	ABQ43627	Oligonucleotide fo
c 230	17	0.5	532	22	AA142831	Probe #11517 used	303	17	0.5	897	20	AAZ11988	Neisseria gonorrhoe
c 231	17	0.5	532	24	AB511017	Human genome-deriv	304	17	0.5	897	24	ABN71052	Streptococcus poly
c 232	17	0.5	540	19	AA114446	H. pylori GHP0 892	305	17	0.5	902	24	ABQ30942	Oligonucleotide fo
c 233	17	0.5	550	24	ABK11646	Human cDNA encodin	306	17	0.5	902	24	ABQ30943	Oligonucleotide fo
c 234	17	0.5	578	22	AAK08520	Human brain expres	c 307	17	0.5	906	22	AAC91351	Human polynucleoti
c 235	17	0.5	578	22	AAK34404	Human bone marrow	c 308	17	0.5	909	21	AAC76460	Human ORFX ORF2015
c 236	17	0.5	578	22	AA140127	Probe #8813 used t	c 309	17	0.5	917	24	ABQ30772	Oligonucleotide fo
c 237	17	0.5	581	23	ABV44584	Human prostate exp	c 310	17	0.5	917	24	ABQ30773	Oligonucleotide fo
c 238	17	0.5	585	23	ABL24939	Drosophila melanog	c 311	17	0.5	921	24	ABQ22926	Oligonucleotide fo
c 239	17	0.5	585	24	ABN60642	Human cancer relat	c 312	17	0.5	921	24	ABQ22927	Oligonucleotide fo
c 240	17	0.5	594	24	ABQ56232	Human ovarian anti	c 313	17	0.5	924	24	ABQ44484	Oligonucleotide fo
c 241	17	0.5	594	24	ABQ32006	Oligonucleotide fo	314	17	0.5	924	24	ABQ44485	Oligonucleotide fo
c 242	17	0.5	594	24	ABQ32007	Oligonucleotide fo	315	17	0.5	925	21	AAA93918	Rice bZIP gene MN3
c 243	17	0.5	595	23	AA87162	DNA encoding novel	316	17	0.5	932	24	ABQ40354	Oligonucleotide fo
c 244	17	0.5	605	22	AAK70580	Human immune/haema	c 317	17	0.5	932	24	ABQ40355	Oligonucleotide fo
c 245	17	0.5	605	22	AAH69784	Human cervical can	c 318	17	0.5	942	23	AA577739	DNA encoding novel
c 246	17	0.5	607	24	ABQ20014	Oligonucleotide fo	319	17	0.5	942	24	ABN69421	Streptococcus poly
c 247	17	0.5	607	24	ABQ20015	Oligonucleotide fo	320	17	0.5	960	24	ABK77876	Bacillus clausii g
c 248	17	0.5	611	22	AAH05991	Human cDNA clone (	c 321	17	0.5	980	22	AA509471	Human DNA containi
c 249	17	0.5	612	22	AAH68969	Human cervical can	c 322	17	0.5	981	18	AA789242	Clone containing h
c 250	17	0.5	640	22	AAH70632	Human cervical can	c 323	17	0.5	981	19	AAV63644	Human telomerase R
c 251	17	0.5	646	21	AAK07854	Fusarium venenatum	c 324	17	0.5	981	19	AAV41194	PstI fragment cont
c 252	17	0.5	651	24	ABK36136	cDNA sequence #527	c 325	17	0.5	981	19	AAV19481	Human telomerase h
c 253	17	0.5	662	24	ABQ14048	Oligonucleotide fo	c 326	17	0.5	981	20	AAZ23262	Human clone 28-1 c
c 254	17	0.5	662	24	ABQ14049	Oligonucleotide fo	c 327	17	0.5	981	21	AAA37564	Human telomerase n
c 255	17	0.5	667	22	AAH70942	Human cervical can	c 328	17	0.5	981	21	AAH45398	Human ADAM10 promo
c 256	17	0.5	680	18	AA758803	Human telomerase e	c 329	17	0.5	981	23	AA515442	PstI fragment cont
c 257	17	0.5	688	21	AAV11899	Aspergillus oryzae	c 330	17	0.5	981	24	AAZ24234	Human telomerase (
c 258	17	0.5	697	21	AAV15055	Trichoderma reesei	c 331	17	0.5	982	24	ABQ36674	Oligonucleotide fo
c 259	17	0.5	710	22	AAH70524	Human cervical can	c 332	17	0.5	982	24	ABQ36675	Oligonucleotide fo
c 260	17	0.5	710	22	AAH78555	DNA encoding novel	c 333	17	0.5	984	22	AA727606	Mevulonate pathway
c 261	17	0.5	713	22	AAH08611	Human cDNA clone (	334	17	0.5	990	22	AA168313	Rice NIF1 related
c 262	17	0.5	716	24	ABQ32384	Oligonucleotide fo	335	17	0.5	1004	21	AAZ44468	Zea mays DNA fragm
c 263	17	0.5	716	24	ABQ32385	Oligonucleotide fo	c 336	17	0.5	1006	21	AAZ97042	Human secreted pro
c 264	17	0.5	724	24	ABQ49084	Oligonucleotide fo	337	17	0.5	1037	23	ABV22330	Human prostate exp
c 265	17	0.5	724	24	ABQ49085	Oligonucleotide fo	c 338	17	0.5	1037	23	ABV23506	Human prostate exp
c 266	17	0.5	741	22	AAK93489	Human cDNA clone r	339	17	0.5	1037	23	ABV28156	Human prostate exp
c 267	17	0.5	746	24	AA561926	Porcine muscular s	340	17	0.5	1037	23	ABV29362	Human prostate exp
c 268	17	0.5	747	22	AAH67861	C glutamicum codin	c 341	17	0.5	1038	22	AAZ19380	Human CCAAT/enhanc
c 269	17	0.5	748	22	AAK67220	Human immune/haema	c 342	17	0.5	1038	22	AA514696	Human cDNA encodin
c 270	17	0.5	748	22	AAK77725	Human immune/haema	c 343	17	0.5	1038	23	ABK51402	DNA encoding human
c 271	17	0.5	762	22	AAH05372	Human cDNA clone (	c 344	17	0.5	1038	24	AAZ44084	Human C/EBP-beta p
c 272	17	0.5	762	22	AAH07885	Human cDNA clone (	c 345	17	0.5	1038	24	ABK51997	DNA encoding wild-
c 273	17	0.5	781	21	AAZ33349	Human secreted pro	c 346	17	0.5	1039	24	ABQ38044	Oligonucleotide fo
c 274	17	0.5	782	24	ABQ18794	Oligonucleotide fo	c 347	17	0.5	1039	24	ABQ38045	Oligonucleotide fo
c 275	17	0.5	782	24	ABQ18795	Oligonucleotide fo	c 348	17	0.5	1042	23	ABK51410	Human CCAAT enhanc
c 276	17	0.5	783	24	ABQ53582	Oligonucleotide fo	c 349	17	0.5	1045	21	AAV13783	Aspergillus oryzae
c 277	17	0.5	783	24	ABQ53583	Oligonucleotide fo	350	17	0.5	1071	22	AAH33017	Human colon cancer
c 278	17	0.5	800	24	ABQ33104	Oligonucleotide fo	c 351	17	0.5	1071	24	ABL89534	Human polynucleoti
c 279	17	0.5	800	24	ABQ33105	Oligonucleotide fo	352	17	0.5	1084	22	AAV31621	Mycobacterium tube
c 280	17	0.5	800	24	ABQ36102	Oligonucleotide fo	353	17	0.5	1119	24	ABQ90245	M. capsulatus gene
c 281	17	0.5	803	24	AA561925	Porcine muscular s	354	17	0.5	1140	22	AAV31648	Mycobacterium tube
c 282	17	0.5	811	24	ABQ38800	Oligonucleotide fo	c 355	17	0.5	1149	21	AAV12939	Aspergillus oryzae
c 283	17	0.5	811	24	ABQ38801	Oligonucleotide fo	356	17	0.5	1168	24	ABQ47030	Oligonucleotide fo
c 284	17	0.5	814	22	AA194797	Human neuroblastom	c 357	17	0.5	1168	24	ABQ47031	Oligonucleotide fo
c 285	17	0.5	821	24	ABQ37728	Oligonucleotide fo	358	17	0.5	1174	22	AAK53200	Human polynucleoti
c 286	17	0.5	821	24	ABQ37729	Oligonucleotide fo	359	17	0.5	1195	24	ABQ31770	Oligonucleotide fo
c 287	17	0.5	826	22	AAH03359	Human cDNA clone (	c 360	17	0.5	1195	24	ABQ31771	Oligonucleotide fo
c 288	17	0.5	852	22	AA194415	Human neuroblastom	c 361	17	0.5	1221	22	AAC85144	Human early growth
c 289	17	0.5	856	22	AAH34310	Human colon cancer	362	17	0.5	1235	24	ABQ32040	Oligonucleotide fo
c 290	17	0.5	856	22	AAZ11977	Neisseria meningit	c 363	17	0.5	1235	24	ABQ32041	Oligonucleotide fo
c 291	17	0.5	859	20	AAZ11977	N. meningitidis Me	364	17	0.5	1242	20	AAZ11978	Neisseria meningit
c 292	17	0.5	859	21	AAZ11977	Human telomerase R	365	17	0.5	1242	20	AAZ11979	Neisseria meningit
c 293	17	0.5	867	20	AAZ07250	DNA encoding novel	366	17	0.5	1246	22	AAZ11980	Neisseria gonorrhoe
c 294	17	0.5	867	23	AA593950	DNA encoding novel	367	17	0.5	1255	19	AAV46408	Corynebacterium gl
c 295	17	0.5	868	24	ABQ40660	Oligonucleotide fo	368	17	0.5	1257	19	AAV46408	Human tumour suppr
c 296	17	0.5	868	24	ABQ40661	Oligonucleotide fo	369	17	0.5	1257	19	AAV46393	Human tumour suppr
c 297	17	0.5	869	23	ABL17159	Oligonucleotide fo	370	17	0.5	1257	20	AAV57674	Human tumour suppr
c 298	17	0.5	869	23	ABL17159	Drosophila melanog	371	17	0.5	1265	22	AAV57674	Human SEC6 DNA seq
c 299	17	0.5	871	18	AAZ93326	Exon 1 of gene of	372	17	0.5	1267	24	ABQ40672	Oligonucleotide fo
c 300	17	0.5	872	22	AA526147	Human cDNA encodin	c 373	17	0.5	1267	24	ABQ40673	Oligonucleotide fo
c 301	17	0.5	890	24	ABQ43626	Oligonucleotide fo	374	17	0.5	1281	9	AAH81257	Mutant ara gene e

375	17	0.5	1281	13	AAQ29507	EPSP synthase wild	c 448	17	0.5	2408	23	ABL14691	Drosophila melanog
376	17	0.5	1281	13	AAQ29508	EPSP synthase muta	c 449	17	0.5	2425	17	AAT10283	Gene for RNA compo
377	17	0.5	1281	15	AAQ64903	EPSP synthase gene	c 450	17	0.5	2425	17	AAT11027	DNA encoding the h
378	17	0.5	1281	15	AAQ64904	Mutant EPSP syntha	c 451	17	0.5	2426	19	AAV22994	DNA containing hum
379	17	0.5	1286	23	AA564240	DNA encoding novel	c 452	17	0.5	2463	22	AAK94536	Human full-length
c 380	17	0.5	1318	22	AAI25841	Murine SHD1 protel	c 453	17	0.5	2467	24	ABQ91979	Human NF-kB activa
c 381	17	0.5	1372	21	AAZ40202	O. sativa nicotian	454	17	0.5	2471	21	AAA73711	Human PITSURE prot
c 382	17	0.5	1372	24	ABQ15720	Oligonucleotide fo	455	17	0.5	2471	24	ABL51630	Human PITSURE prot
c 383	17	0.5	1372	24	ABQ15721	Oligonucleotide fo	456	17	0.5	2493	23	AA888122	DNA encoding novel
c 384	17	0.5	1375	24	ABQ25200	Oligonucleotide fo	c 457	17	0.5	2585	23	ABL24938	Drosophila melanog
c 385	17	0.5	1375	24	ABQ25201	Oligonucleotide fo	c 458	17	0.5	2588	21	AA876756	Human ORFX ORF2311
c 386	17	0.5	1398	23	AA567887	DNA encoding novel	459	17	0.5	2595	22	AAH33385	Human colon cancer
c 387	17	0.5	1408	24	AA562579	CDNA sequence #366	460	17	0.5	2617	21	AAH18171	Lung cancer associ
c 388	17	0.5	1440	20	AAV64277	Human EGR-2 cDNA.	c 461	17	0.5	2653	23	ABL19844	Drosophila melanog
c 389	17	0.5	1460	23	AAI20291	Drosophila melanog	462	17	0.5	2672	16	AAT08693	Cephamycin biosynt
390	17	0.5	1482	21	AA44746	Zea mays DNA fragm	463	17	0.5	2703	23	ABL11233	Drosophila melanog
391	17	0.5	1519	21	AA476492	Human ORFX ORF2047	464	17	0.5	2717	19	AAT96741	DNA coding K. oxyt
392	17	0.5	1521	23	ABL17483	Drosophila melanog	465	17	0.5	2723	23	ABL27612	Drosophila melanog
393	17	0.5	1545	23	AA554038	Pseudomonas aerugi	c 466	17	0.5	2745	23	ABL22957	Drosophila melanog
394	17	0.5	1548	22	AAT68317	Rice NIF1 related	467	17	0.5	2748	23	ABL09795	Drosophila melanog
395	17	0.5	1574	23	AA575663	DNA encoding novel	c 468	17	0.5	2759	20	AAH28295	Human CYP3A4 gene.
396	17	0.5	1607	23	ABL20289	Drosophila melanog	469	17	0.5	2786	22	AAH34862	Human colon cancer
397	17	0.5	1610	24	ABK09767	Human ovarian tumo	470	17	0.5	2786	24	ABL90450	Human polynucleoti
c 398	17	0.5	1629	18	AAI66504	Gene for rice mito	471	17	0.5	2796	19	AAH22648	Homo sapiens trans
399	17	0.5	1664	24	ABA99677	Murine osteoporosi	472	17	0.5	2812	22	AAH17702	Human cDNA sequenc
400	17	0.5	1725	19	AAV45137	Human lysyl-oxidas	c 473	17	0.5	2813	15	AAQ73464	Human Egr-2. Homo
401	17	0.5	1725	24	ABA95144	Human regulatory p	c 474	17	0.5	2817	10	AAH91816	Human Egr-2. Homo
402	17	0.5	1756	23	AA585527	Human prostate exp	475	17	0.5	2829	22	AAK51576	Human polynucleoti
403	17	0.5	1701	23	ABV25058	DNA encoding novel	476	17	0.5	2833	22	AAH23416	Human SEC7 DNA seq
c 404	17	0.5	1713	23	AA591801	Human cDNA sequenc	477	17	0.5	2835	22	AAK52560	Human polynucleoti
405	17	0.5	1720	22	AAH16989	Homo sapiens P-TEN	c 478	17	0.5	2849	24	ABN95612	Gene #2110 used to
c 406	17	0.5	1725	19	AAV45137	Human lysyl-oxidas	c 479	17	0.5	2887	24	AAK99403	DNA of APP related
407	17	0.5	1756	23	AA585527	DNA encoding novel	c 480	17	0.5	2910	24	ABK84640	Human cDNA differe
c 408	17	0.5	1765	20	AAZ07247	Human telomerase R	c 481	17	0.5	2910	24	ABN97289	Gene #3787 used to
c 409	17	0.5	1766	24	AB199779	Mouse ischaemic co	c 482	17	0.5	2910	24	AA561884	Lung small cell ca
c 410	17	0.5	1773	18	AAT93354	Gene of IMAGE clon	c 483	17	0.5	2927	22	AAK86823	Human immune/haema
c 411	17	0.5	1773	21	AAZ29548	Zea mays RecA-like	484	17	0.5	2929	24	AAH31710	Rat SNF1/AMPK-Reta
412	17	0.5	1776	24	ABK84225	Human cDNA differe	485	17	0.5	2930	24	ABQ88181	Human osteoblast d
413	17	0.5	1776	24	AA517569	DNA encoding human	486	17	0.5	2930	24	ABK70285	Human lung cancer
414	17	0.5	1784	20	AAK07339	Dual specificity p	c 487	17	0.5	2934	24	AB199306	Mouse ischaemic co
415	17	0.5	1791	16	AAI04154	BCG DapB gene. My	c 488	17	0.5	2980	24	AA594915	Human DNA sequenc
c 416	17	0.5	1796	23	ABL09605	Drosophila melanog	489	17	0.5	3044	24	ABQ88180	Human osteoblast d
c 417	17	0.5	1850	24	ABQ69203	Listeria monocytog	490	17	0.5	3044	24	ABR84552	Human cDNA differe
c 418	17	0.5	1910	21	AAE21187	Human low adenosin	491	17	0.5	3044	24	ABN97223	Gene #3721 used to
c 419	17	0.5	1910	21	AAA35065	Human adenosine re	492	17	0.5	3047	24	ABK09792	Human ovarian tumo
c 420	17	0.5	1910	22	AAI19379	Human CCAAT/enhanc	493	17	0.5	3072	24	ABQ88182	Human osteoblast d
c 421	17	0.5	1910	23	ABK51401	Human CCAAT enhanc	494	17	0.5	3115	21	AAK98113	Human colon cancer
c 422	17	0.5	1910	24	AAI44083	Human C/EBP-beta p	495	17	0.5	3160	23	AAK46332	Human tumour suppr
c 423	17	0.5	1910	24	ABK51996	DNA encoding wild-	496	17	0.5	3160	20	AAK57671	Human tumour suppr
c 424	17	0.5	1910	24	ABK84563	Human cDNA differe	497	17	0.5	3160	24	ABR84480	Human cDNA differe
c 425	17	0.5	1910	24	ABL94237	Human C/EBP beta-e	498	17	0.5	3160	24	ABK59528	Human phosphatase
c 426	17	0.5	1913	12	AAQ15015	Nuclear factor C/E	c 499	17	0.5	3169	22	AAI58822	Human polynucleoti
c 427	17	0.5	1962	20	AAK57672	Mouse tumour suppr	500	17	0.5	3184	22	ABA09269	Human IGFALS homol
428	17	0.5	1967	24	ABN95920	Gene #2418 used to	501	17	0.5	3203	22	AAI60608	Human polynucleoti
429	17	0.5	2004	12	AAQ13359	Human mevalonate k	502	17	0.5	3217	12	AAQ12002	Full-length rat an
430	17	0.5	2071	22	AAH24555	zinc finger protei	c 503	17	0.5	3317	23	ABV24474	Human prostate exp
c 431	17	0.5	2007	22	AAH16766	Human cDNA sequenc	504	17	0.5	3330	16	AAK86857	Human very low den
c 432	17	0.5	2016	22	AA526565	Human cDNA encodin	505	17	0.5	3349	20	AAK60577	Human disintegrin
c 433	17	0.5	2022	22	AAH15029	Human cDNA sequenc	506	17	0.5	3402	24	AAH35330	Human toll like re
c 434	17	0.5	2037	22	AAH51958	Mycobacterium tube	507	17	0.5	3410	22	AAH45399	Human ADAM10 codin
435	17	0.5	2047	22	AAK52216	Human polynucleoti	508	17	0.5	3410	22	AAH83470	Human ADAM10 DNA.
436	17	0.5	2067	24	ABL34301	Human immune syste	c 509	17	0.5	3460	22	AAH72627	Human cervical can
437	17	0.5	2067	24	ABK28460	DNA transcription	c 510	17	0.5	3460	22	AAH72735	Human immune syste
c 438	17	0.5	2088	22	AAH14076	Human cDNA sequenc	c 511	17	0.5	3504	24	ABL34314	Drosophila melanog
439	17	0.5	2139	19	AAV56670	Human ELL2 cDNA.	512	17	0.5	3523	23	ABL09879	Human immune syste
440	17	0.5	2160	20	AAK57700	Mouse MMAC1 gene.	c 513	17	0.5	3531	21	AAK58368	Human PRO290 nucle
c 441	17	0.5	2171	21	AAE21188	Human low adenosin	c 514	17	0.5	3531	21	AAK64934	Membrane-bound pro
c 442	17	0.5	2171	21	AAK535066	Human adenosine re	c 515	17	0.5	3531	22	AAK44100	Human PRO290 (UNQ2
443	17	0.5	2238	23	ABL28399	Drosophila melanog	c 516	17	0.5	3546	22	AAE72802	Secreted protein 9
444	17	0.5	2274	23	AAI28784	Human nuclear horm	517	17	0.5	3597	23	ABL09813	Drosophila melanog
445	17	0.5	2305	24	ABA94651	Human secreted red	518	17	0.5	3621	23	ABL17482	Drosophila melanog
c 446	17	0.5	2328	24	AAI67786	Human lysyl oxidas	519	17	0.5	3656	17	AAK36751	VLDL receptor gene
c 447	17	0.5	2406	17	AAI33677	Human transcriptio	520	17	0.5	3773	23	ABL11845	Drosophila melanog

c 521	17	0.5	3782	23	ABL28681	Drosophila melanog	594	17	0.5	8801	24	ABL33740	Human immune syste
c 522	17	0.5	3969	24	ABN95715	Gene #2213 used to	595	17	0.5	8801	24	ABK28285	DNA transcription
c 523	17	0.5	3978	23	AAS87429	DNA encoding novel	c 596	17	0.5	8841	22	AAK78951	Human immune/haema
c 524	17	0.5	4010	16	AAO84303	Human NF-AT120, X	c 597	17	0.5	8841	22	AAK78952	Human immune/haema
c 525	17	0.5	4138	24	AAI72578	Human Na+-driven C	c 598	17	0.5	8919	23	ABL20699	Drosophila melanog
c 526	17	0.5	4159	23	ABL10310	Drosophila melanog	c 599	17	0.5	8968	23	ABL11297	Drosophila melanog
c 527	17	0.5	4171	22	AAH18676	Human cDNA sequenc	c 600	17	0.5	9002	23	ABL09604	Drosophila melanog
c 528	17	0.5	4188	22	AAF98721	Human late stage o	c 601	17	0.5	9072	18	AAV74356	Staphylococcus aur
c 529	17	0.5	4250	24	ABK86098	Human cDNA for hae	c 602	17	0.5	9163	24	ABL45808	Human endothelial
c 530	17	0.5	4311	23	ABL01926	Drosophila melanog	c 603	17	0.5	9592	17	AAAT36752	Adenovirus vector
c 531	17	0.5	4321	24	AB199593	Mouse ischaemic co	c 604	17	0.5	9990	22	AAH26123	Mouse prion protei
c 532	17	0.5	4351	18	AAAT97850	M. halophilus ecto	c 605	17	0.5	9990	22	AAH26136	Mouse PrP/human ta
c 533	17	0.5	4373	23	ABL28398	Drosophila melanog	c 606	17	0.5	10250	24	ABN80082	Human chemically m
c 534	17	0.5	4389	23	ABL14960	Drosophila melanog	c 607	17	0.5	10336	22	ABA21257	Human nervous syst
c 535	17	0.5	4389	23	ABL20288	Drosophila melanog	c 608	17	0.5	10394	23	ABL11060	Drosophila melanog
c 536	17	0.5	4391	23	ABL20290	Drosophila melanog	c 609	17	0.5	10682	24	ABN80085	Human chemically m
c 537	17	0.5	4402	24	ABK12867	Drosophila melanog	c 610	17	0.5	10747	21	AAK60207	Sequence used to p
c 538	17	0.5	4445	22	AAH18522	Human cDNA sequenc	c 611	17	0.5	10747	21	AAA52319	Human keratin K1 (
c 539	17	0.5	4534	20	AAK24069	Human 53BP2 DNA.	c 612	17	0.5	11049	24	ABL92218	Chemically treated
c 540	17	0.5	4534	21	AAA09324	Human cancer assoc	c 613	17	0.5	11049	24	ABL49321	Human polynucleoti
c 541	17	0.5	4641	21	AAF16251	Human prostate can	c 614	17	0.5	11049	24	ABL32668	Human immune syste
c 542	17	0.5	4751	22	AAK26783	Human genomic DNA	c 615	17	0.5	11500	23	ABL28680	Drosophila melanog
c 543	17	0.5	4753	23	ABL15906	Drosophila melanog	c 616	17	0.5	11888	22	AAI07198	Human reproductive
c 544	17	0.5	4957	24	ABN95656	Gene #2154 used to	c 617	17	0.5	12409	24	AAK63312	Chemically pretrea
c 545	17	0.5	4957	24	ABL68507	Kidney cancer rela	c 618	17	0.5	12550	23	ABL29300	Drosophila melanog
c 546	17	0.5	4957	24	ABL68869	Kidney cancer rela	c 619	17	0.5	12870	23	ABL114952	Drosophila melanog
c 547	17	0.5	4980	23	ABL11232	Drosophila melanog	c 620	17	0.5	13321	22	AAK46421	Tumour suppressor
c 548	17	0.5	5045	23	ABL09794	Drosophila melanog	c 621	17	0.5	13322	23	ABL09812	Drosophila melanog
c 549	17	0.5	5128	21	AAK75835	DNA encoding the h	c 622	17	0.5	13960	23	AAK87164	DNA encoding novel
c 550	17	0.5	5310	24	ABL70417	Chemically treated	c 623	17	0.5	14093	23	ABK42857	Genomic sequence #
c 551	17	0.5	5310	24	ABK31444	Signal transductio	c 624	17	0.5	14170	23	ABL11044	Drosophila melanog
c 552	17	0.5	5378	24	ABL33684	Human immune syste	c 625	17	0.5	14372	23	ABL09878	Drosophila melanog
c 553	17	0.5	5385	24	AAI72577	Murine Na+-driven	c 626	17	0.5	14708	23	ABL13296	Drosophila melanog
c 554	17	0.5	5404	23	ABL14953	Drosophila melanog	c 627	17	0.5	16818	22	ABA15470	Human nervous syst
c 555	17	0.5	5467	24	ABL34109	Human immune syste	c 628	17	0.5	16818	22	ABA16242	Human nervous syst
c 556	17	0.5	5499	24	ABQ66972	Human angiogenesis	c 629	17	0.5	16818	22	ABA18914	Human nervous syst
c 557	17	0.5	5631	23	ABL17158	Drosophila melanog	c 630	17	0.5	16842	22	AAK46411	Tumour suppressor
c 558	17	0.5	5749	23	ABL22956	Drosophila melanog	c 631	17	0.5	16842	24	ABL70383	Chemically treated
c 559	17	0.5	5798	24	ABQ67012	Drosophila melanog	c 632	17	0.5	16842	24	AAK61335	Human gene regulat
c 560	17	0.5	5798	24	ABL32169	Human immune syste	c 633	17	0.5	16842	24	ABK31418	Signal transductio
c 561	17	0.5	5814	24	ABN59737	Novel human coding	c 634	17	0.5	17528	24	ABL34600	Human immune syste
c 562	17	0.5	5887	24	ABQ71051	Listeria monocytog	c 635	17	0.5	17893	24	ABL33364	Human immune syste
c 563	17	0.5	5997	24	ABL70285	Chemically treated	c 636	17	0.5	17951	23	ABL11844	Drosophila melanog
c 564	17	0.5	5997	24	AAK61210	Human gene regulat	c 637	17	0.5	18683	24	ABL54333	Chemically treated
c 565	17	0.5	5997	24	ABK31310	Signal transductio	c 638	17	0.5	18683	23	ABL32312	Human immune syste
c 566	17	0.5	6065	24	ABK70579	Chemically treated	c 639	17	0.5	18871	23	ABK42430	Genomic sequence #
c 567	17	0.5	6065	24	AAK61260	Human gene regulat	c 640	17	0.5	19233	24	ABL45345	Human polynucleoti
c 568	17	0.5	6065	24	ABK31356	Signal transductio	c 641	17	0.5	20486	23	ABL20698	Drosophila melanog
c 569	17	0.5	6223	23	ABL14690	Drosophila melanog	c 642	17	0.5	22275	23	ABL25566	Drosophila melanog
c 570	17	0.5	6307	24	ABL33341	Human immune syste	c 643	17	0.5	24978	20	AAK60209	SEQ ID 3 of US5914
c 571	17	0.5	6334	22	AAK46458	Tumour suppressor	c 644	17	0.5	24979	21	AAA52321	Genomic DNA sequen
c 572	17	0.5	6375	24	ABL58984	AAV expression vec	c 645	17	0.5	26657	22	AAI72317	Human transporter
c 573	17	0.5	6599	24	ABN80183	Human chemically m	c 646	17	0.5	26997	22	AAK46748	Tumour suppressor
c 574	17	0.5	6731	24	ABL33061	Human immune syste	c 647	17	0.5	28730	21	AAA81486	N. meningitidis pa
c 575	17	0.5	6734	24	ABK92180	Prostate cancer-as	c 648	17	0.5	32134	22	AAI99172	Human excretory re
c 576	17	0.5	6954	24	ABK33390	Human immune syste	c 649	17	0.5	32134	22	AAI63522	Human kidney relat
c 577	17	0.5	6977	22	AAK45627	Tumour suppressor	c 650	17	0.5	32192	22	AAI99173	Human excretory re
c 578	17	0.5	7503	24	ABQ67009	Human angiogenesis	c 651	17	0.5	32192	22	AAI63523	Human kidney relat
c 579	17	0.5	7503	24	ABL33548	Human immune syste	c 652	17	0.5	35100	22	AAK65700	Human immune/haema
c 580	17	0.5	7511	24	ABL33283	Human immune syste	c 653	17	0.5	35100	22	AAK69767	Human immune/haema
c 581	17	0.5	7625	22	AAI04533	Human reproductive	c 654	17	0.5	35115	22	AAK65699	Human immune/haema
c 582	17	0.5	7625	23	ABL97456	Human testicular a	c 655	17	0.5	35115	22	AAK69766	Human immune/haema
c 583	17	0.5	7633	22	AAI04536	Human reproductive	c 656	17	0.5	35465	22	AAK54723	Nucleotide sequenc
c 584	17	0.5	7633	23	ABL97459	Human testicular a	c 657	17	0.5	37856	21	AAAI1992	S. cellulosum DNA
c 585	17	0.5	7667	22	AAK46334	Tumour suppressor	c 658	17	0.5	37973	24	ABL34197	Human immune syste
c 586	17	0.5	7669	22	AAK46597	Tumour suppressor	c 659	17	0.5	48037	22	AAK84729	Human immune/haema
c 587	17	0.5	7669	24	ABL33846	Human immune syste	c 660	17	0.5	48037	22	AAK85983	Human immune/haema
c 588	17	0.5	7900	22	AAK46759	Tumour suppressor	c 661	17	0.5	48045	22	AAK84730	Human immune/haema
c 589	17	0.5	8093	21	AAK81512	N. meningitidis pa	c 662	17	0.5	48045	22	AAK85984	Human immune/haema
c 590	17	0.5	8451	23	ABL11062	Drosophila melanog	c 663	17	0.5	56485	21	AAK84476	N. meningitidis pa
c 591	17	0.5	8711	22	AAK26809	Human genomic DNA	c 664	17	0.5	63563	22	AAK28546	Genomic fragment #
c 592	17	0.5	8771	24	ABL33825	Human immune syste	c 665	17	0.5	72750	21	AAA81468	N. meningitidis pa
c 593	17	0.5	8801	22	AAK45436	Chemically pretrea	c 666	17	0.5	73947	23	ABL07230	Drosophila melanog

c 667	17	0.5	83391	24	ABQ67094	Human angio genesis	740	0.5	210	22	ABA36815	Probe #15281 for g
668	17	0.5	83946	24	ABQ88101	Human osteoblast d	741	0.5	210	22	AAK18195	Human brain expres
669	17	0.5	349980	21	AAF21544	Neisseria meningit	742	0.5	210	22	AAK44089	Human bone marrow
c 670	17	0.5	349980	21	AAF21544	Neisseria meningit	743	0.5	210	22	AAI24723	Probe #14656 for g
c 671	17	0.5	349980	21	AAF21609	Neisseria meningit	744	0.5	210	22	AAI50098	Probe #18784 used
c 672	17	0.5	349980	21	AAF21610	Neisseria meningit	745	0.5	210	22	ABSI18323	Human genome-deriv
c 673	17	0.5	349980	22	AAH41223	Pyrococcus abyssii	c 746	0.5	216	24	ABN25051	Human ORFX polynuc
c 674	17	0.5	349980	22	AAH41224	Pyrococcus abyssii	747	0.5	217	24	ABN25563	Human ORFX polynuc
c 675	17	0.5	349980	22	AAH41225	Pyrococcus abyssii	748	0.5	225	21	AAI19040	Human gene signatu
c 676	17	0.5	349980	22	AAH68533	C glutamicum codin	c 749	0.5	229	16	AAI22982	Human secreted pro
c 677	17	0.5	368004	24	ABL57909	Human transporter	750	0.5	229	21	AAI238342	Human secreted pro
c 678	17	0.5	837096	21	AAAB1489	N. meningitidis pa	c 751	0.5	230	22	ABA73231	Human foetal liver
c 679	17	0.5	1437668	21	AAAB1490	N. meningitidis B	c 752	0.5	230	22	ABA38655	Probe #17121 for g
680	17	0.5	1664976	19	AAV21209	Methanococcus jann	c 753	0.5	230	22	AAK21666	Human brain expres
681	17	0.5	2365599	24	ABA90521	Genomic sequence o	c 754	0.5	230	22	AAK47831	Human bone marrow
682	17	0.5	4403765	22	AAI199683	Mycobacterium tube	c 755	0.5	230	22	AAI26055	Probe #15988 for g
c 683	17	0.5	4403765	22	AAI199683	Mycobacterium tube	c 756	0.5	230	22	AAI53659	Probe #22345 used
c 684	17	0.5	4411529	22	AAI199682	Mycobacterium tube	c 757	0.5	230	24	ABS21830	Human genome-deriv
c 685	17	0.5	4411529	22	AAI199682	Mycobacterium tube	758	0.5	235	21	AAI30438	Human secreted pro
c 686	16	0.5	18	20	AAZ31863	Human G-alpha-13 a	c 759	0.5	243	22	AAI22982	Human lung antigen
687	16	0.5	20	22	AAI191298	Human E2F transcri	c 760	0.5	250	21	AAI18381	Human secreted pro
688	16	0.5	25	19	AAV07946	Helicobacter pylor	c 761	0.5	256	20	AAV88334	Human secreted pro
689	16	0.5	26	19	AAV07943	Helicobacter pylor	762	0.5	258	21	AAI68906	EST clone FX194.
c 690	16	0.5	26	21	AAZ55457	Neisseria species	c 763	0.5	263	21	AAI31075	Rat ADIPO7 coding
c 691	16	0.5	26	22	AAZ44219	Neisseria meningit	c 764	0.5	267	22	AAI28361	Plant microsatelli
c 692	16	0.5	26	22	AAZ44241	Neisseria meningit	765	0.5	271	21	AAI28361	Genomic sequence #
c 693	16	0.5	26	22	AAZ44243	Neisseria meningit	c 766	0.5	271	22	AAI28361	Human secreted pro
c 694	16	0.5	26	22	AAZ44245	Neisseria meningit	c 767	0.5	274	24	ABL71358	Novel cDNA encodin
c 695	16	0.5	26	22	AAZ44247	Neisseria meningit	768	0.5	276	24	ABL71358	Corn tassell-derive
c 696	16	0.5	26	22	AAZ44344	Neisseria meningit	769	0.5	283	24	ABK62781	Corn tassell-derive
c 697	16	0.5	26	22	AAI17062	Neisseria meningit	c 770	0.5	284	13	AAQ22701	Rat sequence diffe
c 698	16	0.5	26	22	AAI17107	N. meningitidis fu	c 771	0.5	284	24	ABN95468	Sequence encoding
699	16	0.5	31	22	AAZ44240	Neisseria meningit	c 772	0.5	284	24	ABK64465	Gene #1966 used to
700	16	0.5	31	22	AAZ44244	Neisseria meningit	c 773	0.5	284	24	ABK64465	Human benign prost
c 701	16	0.5	35	22	AAZ44249	Neisseria meningit	774	0.5	285	24	ABN21304	Kidney cancer rela
702	16	0.5	47	21	AAZ68103	Human map-related	775	0.5	287	22	AAI11882	Human ORFX polynuc
c 703	16	0.5	47	24	AAZ47108	Peptide presentati	776	0.5	287	22	AAI20772	Human breast cance
c 704	16	0.5	47	24	AAZ47109	Peptide presentati	c 777	0.5	287	22	AAI54240	Murine translation
c 705	16	0.5	50	21	AAI26760	Human secreted pro	c 778	0.5	287	24	ABL37093	Human colon tumour
706	16	0.5	50	22	AAI26760	Human secreted pro	779	0.5	289	22	AAI20254	Human breast cance
707	16	0.5	51	22	AAI28808	Oryctolagus cunicu	c 780	0.5	289	24	ABN62942	Human cancer relat
708	16	0.5	54	21	AAI73936	GFP Gly(GGC)5 forw	c 781	0.5	290	24	ABK31117	Plant dwarfing/stu
c 709	16	0.5	54	21	AAI73935	GFP Pro(CCG)5 forw	c 782	0.5	292	24	ABN77115	Human ORF2062 cDNA
710	16	0.5	57	20	AAI60226	Fragment of the PA	783	0.5	298	20	AAI40983	Human secreted pro
c 711	16	0.5	60	24	ABN34980	Human spliced tran	c 784	0.5	300	20	AAI14673	Human gene express
c 712	16	0.5	60	24	ABN42530	Human spliced tran	785	0.5	300	21	AAI1470	Human colon cancer
c 713	16	0.5	65	24	ABN56503	Mouse spliced tran	c 786	0.5	301	21	AAI06635	Human secreted pro
c 714	16	0.5	67	16	AAQ79642	Nucleotide sequenc	787	0.5	301	24	ABK79104	Bacillus Clausii g
715	16	0.5	99	22	ABA47799	Human breast cell	c 788	0.5	304	24	ABL71158	Corn tassell-derive
716	16	0.5	99	22	AAI06347	Probe #6338 used t	c 789	0.5	307	21	AAI40720	Zea mays DNA fragm
717	16	0.5	99	24	ABSI3932	Human genome-deriv	790	0.5	310	24	ABL81208	Human ovarian canc
c 718	16	0.5	105	21	AAI18533	Human secreted pro	791	0.5	313	22	AAI39057	Novel human diagno
719	16	0.5	119	19	AAI16155	Microsatellite mar	c 792	0.5	314	20	AAI31890	Mouse Dix5 5'untra
720	16	0.5	126	21	AAI31060	Plant microsatelli	c 793	0.5	314	20	AAV89554	EST clone CR1190.
721	16	0.5	145	17	AAI28159	Senescence-related	c 794	0.5	314	24	ABS20910	Human genome-deriv
c 722	16	0.5	146	24	ABSI7316	Human genome-deriv	c 795	0.5	315	24	ABS20910	Human colon cancer
723	16	0.5	152	19	AAI16158	Microsatellite mar	c 796	0.5	318	21	AAI31182	Plant microsatelli
c 724	16	0.5	152	21	AAH51644	Sbgl1 exon sequence	797	0.5	321	22	AAI13973	Human 5' UTR of FM
c 725	16	0.5	152	21	AAH51667	Sbgl1 exon sequence	c 798	0.5	322	18	AAI75508	P. americanus anti
c 726	16	0.5	152	21	AAH51679	Sbgl1 exon sequence	c 799	0.5	322	21	AAZ49961	Winter flounder li
c 727	16	0.5	152	21	AAH51679	Sbgl1 exon sequence	c 800	0.5	323	21	AAZ49961	Human secreted pro
728	16	0.5	152	21	AAH51679	Sbgl1 exon sequence	c 801	0.5	330	24	ABN20027	Human ORFX polynuc
729	16	0.5	159	22	AAI68452	Chicken beta-actin	c 802	0.5	330	22	AAI83366	Human polynucleoti
730	16	0.5	160	22	AAI68452	Human immune/haema	c 803	0.5	332	22	AAI83366	Human stresscopin
731	16	0.5	166	20	AAI01048	Septoria sequence	c 804	0.5	339	24	AAI43490	Human brain expres
c 731	16	0.5	171	21	AAI31044	Plant microsatelli	c 805	0.5	342	14	AAQ61118	Human stresscopin
732	16	0.5	179	22	AAI19730	Human breast cance	c 806	0.5	345	24	ABL88684	Conus striatus I-s
733	16	0.5	186	21	AAI33994	Human secreted pro	c 807	0.5	348	22	AAI07101	DNA encoding Group
734	16	0.5	192	21	AAI18555	N. meningitidis pa	c 808	0.5	350	21	AAI56190	Eucalyptus grandis
735	16	0.5	195	21	AAI31565	Plant microsatelli	c 809	0.5	351	23	AAI56190	cDNA #389 encoding
736	16	0.5	200	21	AAI325887	Human secreted pro	c 810	0.5	352	23	AAI56190	Human prostate exp
c 737	16	0.5	204	23	ABL22711	Drosophila melanog	c 811	0.5	355	22	AAI29472	Colon tumour relat
738	16	0.5	208	21	AAI22643	Human secreted pro	c 812	0.5	359	21	AAI29472	Eucalyptus grandis
739	16	0.5	210	22	ABA69986	Human foetal liver	16	0.5	362	24	ABN18066	Human ORFX polynuc

c 813	16	0.5	365	16	AAT19524	Human gene signatu	886	16	0.5	390	17	AAT41957	Modified major tic
c 814	16	0.5	367	22	ABA47063	Human breast cell	887	16	0.5	390	17	AAT41949	Major tick allerge
c 815	16	0.5	367	22	ABA64946	Human foetal liver	888	16	0.5	390	17	AAT41950	Major tick allerge
c 816	16	0.5	367	22	ABA32055	Probe #10521 for g	889	16	0.5	390	17	AAT41951	Modified major tic
c 817	16	0.5	367	22	AK13373	Human brain expres	890	16	0.5	390	17	AAT41952	Modified major tic
c 818	16	0.5	367	22	AK39110	Human bone marrow	891	16	0.5	390	17	AAT41958	Modified major tic
c 819	16	0.5	367	22	AAT19919	Probe #9852 for ge	892	16	0.5	390	17	AAT41959	Modified major tic
c 820	16	0.5	367	22	AAT45114	Probe #13800 used	893	16	0.5	390	17	AAT41960	Modified major tic
c 821	16	0.5	367	22	AAT05633	Probe #5624 used t	894	16	0.5	390	22	AAT89540	Human polynucleoti
c 822	16	0.5	367	22	ABS13192	Human genome-deriv	895	16	0.5	393	15	AQ54064	Der f II Tick majo
c 823	16	0.5	367	24	ABN20251	Human OREF polynuc	896	16	0.5	393	22	AAF67438	Novel human polynu
c 824	16	0.5	369	24	ABN76446	Human OREF1393 cDNA	897	16	0.5	396	22	AAI81388	Human polynucleoti
c 825	16	0.5	372	20	AAV73008	Human adult placen	898	16	0.5	396	22	AAF94816	Human ovarian carc
c 826	16	0.5	372	23	AA568870	DNA encoding novel	c 899	16	0.5	396	22	AAF94872	Human ovarian carc
c 827	16	0.5	372	24	ABQ92033	Human polynucleoti	c 900	16	0.5	396	22	AAF94873	Human ovarian carc
c 828	16	0.5	373	21	AAC22865	Human secreted pro	c 901	16	0.5	396	24	ABT03083	Human ovarian carc
c 829	16	0.5	373	22	ABA58544	Human foetal liver	c 902	16	0.5	396	24	ABT03139	Human ovarian carc
c 830	16	0.5	373	22	AAK06663	Human brain expres	c 903	16	0.5	396	24	ABT03140	Human ovarian carc
c 831	16	0.5	373	22	AAK32359	Human bone marrow	c 904	16	0.5	396	24	ABL48766	Human ovarian carc
c 832	16	0.5	373	22	AAI38206	Probe #6892 used t	c 905	16	0.5	396	24	ABL48822	Ovarian carcinoma
c 833	16	0.5	381	24	ABN21464	Human genome-deriv	c 906	16	0.5	396	24	ABL48823	Ovarian carcinoma
c 834	16	0.5	378	24	ABN24549	Human OREF polynuc	c 907	16	0.5	399	24	ABQ54857	Human ovarian anti
c 835	16	0.5	378	24	ABL81730	Human ovarian carc	c 908	16	0.5	399	24	ABN73862	Human ovarian anti
c 836	16	0.5	379	21	AAA31636	Plant microsatelli	c 909	16	0.5	399	24	AAV78347	Bovine embryonic g
c 837	16	0.5	381	23	ABL41662	Nucleotide sequenc	910	16	0.5	400	18	AAV78347	Human lococortin-re
c 838	16	0.5	381	24	AAK40311	Human OREF polynuc	911	16	0.5	401	22	AAK96048	Staphylococcus aur
c 839	16	0.5	384	20	AAK40311	Human secreted pro	912	16	0.5	401	22	AAK96049	Human neurogulin g
c 840	16	0.5	384	22	AAI81797	Human polynucleoti	913	16	0.5	401	22	AAK97541	Human neurogulin g
c 841	16	0.5	385	21	AAC02615	Human secreted pro	914	16	0.5	401	22	AAK97542	Human neurogulin g
c 842	16	0.5	385	23	AA569852	DNA encoding novel	915	16	0.5	401	22	AAH99146	Human neurogulin g
c 843	16	0.5	386	18	AAAT84649	RFHV glycoprotein	916	16	0.5	402	23	AAK23448	Human neurogulin g
c 844	16	0.5	387	22	AAI29173	Colon tumour relat	917	16	0.5	402	23	AAC09406	Human EST-derived
c 845	16	0.5	387	22	AAK66682	Novel human polynu	918	16	0.5	404	22	ABH42798	Human lung tumour-
c 846	16	0.5	387	22	AAF67407	Novel human polynu	919	16	0.5	406	22	ABA45369	Human secreted pro
c 847	16	0.5	390	15	AAQ73668	Tick allergen Derf	c 920	16	0.5	406	22	ABA55856	Nucleotide sequenc
c 848	16	0.5	390	15	AAQ73671	Modified major tic	c 921	16	0.5	406	22	ABA55856	Human breast cell
c 849	16	0.5	390	15	AAQ73672	Modified major tic	c 922	16	0.5	406	22	ABA25531	Human foetal liver
c 850	16	0.5	390	15	AAQ73670	Tick allergen Derf	c 923	16	0.5	406	22	AAK04075	Probe #3997 for ge
c 851	16	0.5	390	16	AAQ89988	Modified Derf II m	c 924	16	0.5	406	22	AAK29558	Human brain expres
c 852	16	0.5	390	16	AAQ89989	Modified Derf II m	c 925	16	0.5	406	22	AAI14130	Human bone marrow
c 853	16	0.5	390	16	AAQ89990	Modified Derf II m	c 926	16	0.5	406	22	AAI35513	Probe #4063 for ge
c 854	16	0.5	390	16	AAQ89991	Modified Derf II m	c 927	16	0.5	406	22	AAI03980	Probe #4199 used t
c 855	16	0.5	390	16	AAQ89992	Modified Derf II m	c 928	16	0.5	406	24	ABS04112	Probe #3971 used t
c 856	16	0.5	390	16	AAQ89993	Modified Derf II m	c 929	16	0.5	409	22	AAI80936	Human genome-deriv
c 857	16	0.5	390	16	AAQ89994	Modified Derf II m	c 930	16	0.5	411	22	AAK66875	Human polynucleoti
c 858	16	0.5	390	16	AAQ89983	Modified Derf II m	c 931	16	0.5	412	22	ABA58817	Novel human polynu
c 859	16	0.5	390	16	AAQ89984	Modified Derf II m	932	16	0.5	412	22	AAK06971	Human foetal liver
c 860	16	0.5	390	16	AAQ89985	Modified Derf II m	933	16	0.5	412	22	AAK32707	Human brain expres
c 861	16	0.5	390	16	AAQ89986	Modified Derf II m	c 934	16	0.5	412	22	AAI38522	Human bone marrow
c 862	16	0.5	390	16	AAQ89987	Modified Derf II m	c 935	16	0.5	412	23	ABV61956	Human prostate exp
c 863	16	0.5	390	16	AAQ89978	Modified Derf II m	936	16	0.5	412	24	ABS07503	Human genome-deriv
c 864	16	0.5	390	16	AAQ89979	Modified Derf II m	937	16	0.5	416	21	AAK79207	Human lung tumour-
c 865	16	0.5	390	16	AAQ89980	Modified Derf II m	938	16	0.5	416	22	AAK23283	Human lung tumour-
c 866	16	0.5	390	16	AAQ89981	Modified Derf II m	c 939	16	0.5	418	22	AAI11355	Human breast cance
c 867	16	0.5	390	16	AAQ89967	Modified Derf II m	c 940	16	0.5	419	21	AAA31136	Human breast cance
c 868	16	0.5	390	16	AAQ89968	Modified Derf II m	941	16	0.5	420	21	AAK41346	Plant microsatelli
c 869	16	0.5	390	16	AAQ89969	Modified Derf II m	942	16	0.5	423	21	AAK41346	ze mays DNA fragm
c 870	16	0.5	390	16	AAQ89970	Modified Derf II m	943	16	0.5	424	22	AAI10150	Human breast cance
c 871	16	0.5	390	16	AAQ89971	Modified Derf II m	c 944	16	0.5	424	23	AAI81908	Human polynucleoti
c 872	16	0.5	390	16	AAQ89972	Modified Derf II m	945	16	0.5	424	24	ABS57348	CDNA #24 encoding
c 873	16	0.5	390	16	AAQ89973	Modified Derf II m	946	16	0.5	424	24	ABL87144	Human ovarian carc
c 874	16	0.5	390	16	AAQ89974	Modified Derf II m	947	16	0.5	425	21	AAK43732	ze mays DNA fragm
c 875	16	0.5	390	16	AAQ89975	Modified Derf II m	948	16	0.5	425	21	AAZ51838	Maize plastid targ
c 876	16	0.5	390	16	AAQ89976	Modified Derf II m	949	16	0.5	428	24	ABK53867	Human head and nec
c 877	16	0.5	390	16	AAQ89977	Modified Derf II m	c 949	16	0.5	430	24	ABK76055	Bacillus lichenifo
c 878	16	0.5	390	16	AAQ89978	Modified Derf II m	c 950	16	0.5	435	21	AAK56506	Eucalyptus grandis
c 879	16	0.5	390	16	AAQ89979	Modified Derf II m	c 951	16	0.5	439	22	AAI06748	Human reproductive
c 880	16	0.5	390	16	AAQ89966	Modified Derf II m	c 952	16	0.5	439	22	AAI62654	Human breast or ov
c 881	16	0.5	390	16	AAQ89963	Modified Derf II m	c 953	16	0.5	440	21	AAA53334	Myrtaceae microsat
c 882	16	0.5	390	16	AAQ89965	Native Derf II maj	954	16	0.5	440	22	AAH06282	Human cDNA clone (
c 883	16	0.5	390	17	AAI41953	Modified major tic	955	16	0.5	441	21	AAK79262	Human lung tumour-
c 884	16	0.5	390	17	AAI41954	Modified major tic	956	16	0.5	441	23	AAK23338	Human lung tumour-
c 885	16	0.5	390	17	AAI41955	Modified major tic	957	16	0.5	445	22	AAK92100	Human cDNA 5'-end
	16	0.5	390	17	AAI41956	Modified major tic	958	16	0.5	445	22	AAK93630	Human cDNA clone r
										447	22	ABA59050	Human foetal liver

```
959 16 0.5 447 22 ABA27862 Probe #6328 for ge
960 16 0.5 447 22 AAK07223 Human brain expres
961 16 0.5 447 22 AAK32979 Human bone marrow
962 16 0.5 447 22 AAI16281 Probe #6214 for ge
963 16 0.5 447 22 AAI38781 Probe #7467 used t
c 964 16 0.5 447 24 ABS04802 Human genome-deriv
c 965 16 0.5 447 24 ABS07787 Human genome-deriv
c 966 16 0.5 448 22 ABA20103 Human nervous syst
c 967 16 0.5 451 22 ABA13679 Human nervous syst
c 968 16 0.5 451 22 RAF87079 Rice isoleucyl-RNA
c 969 16 0.5 451 24 ABL79824 Human ovarian can
c 970 16 0.5 452 24 ABN75862 Human synthase-lik
c 971 16 0.5 456 21 AAC01184 Human secreted pro
c 972 16 0.5 456 22 ABA42594 Human breast cell
c 973 16 0.5 456 22 ABA53024 Human foetal liver
c 974 16 0.5 456 22 ABA22798 Probe #1264 for ge
c 975 16 0.5 456 22 AAK01272 Human brain expres
c 976 16 0.5 456 22 AAK26733 Human bone marrow
c 977 16 0.5 456 22 AAI11359 Probe #1292 for ge
c 978 16 0.5 456 22 AAI32628 Probe #1314 used t
c 979 16 0.5 456 22 AAI01275 Probe #1266 used t
c 980 16 0.5 456 24 ABS01327 Human genome-deriv
c 981 16 0.5 457 16 AAQ95181 Simple tandem repe
c 982 16 0.5 457 21 AAC01875 Human secreted pro
c 983 16 0.5 458 22 AAK68444 Human immune/haema
c 984 16 0.5 458 22 AAK84980 Human immune/haema
c 985 16 0.5 461 21 AAC16175 Human secreted pro
c 986 16 0.5 462 24 ABK79828 Bacillus clausii g
c 987 16 0.5 464 21 ABA31842 Plant microsattel
c 988 16 0.5 464 22 ABA57359 Human foetal liver
c 989 16 0.5 464 22 ABA26895 Probe #5361 for ge
c 990 16 0.5 464 22 AAK05395 Human brain expres
c 991 16 0.5 464 22 AAK30992 Human bone marrow
c 992 16 0.5 464 22 AAI15496 Probe #5429 for ge
c 993 16 0.5 464 22 AAI36903 Probe #5589 used t
c 994 16 0.5 464 24 AAH06056 Human cdna clone (
c 995 16 0.5 464 22 ABS05741 Human genome-deriv
c 996 16 0.5 465 24 ABL01416 Murine apoptosis r
c 997 16 0.5 468 21 AAC38306 Zea mays DNA fragm
c 998 16 0.5 469 22 AAI33088 Probe #1774 used t
c 999 16 0.5 469 24 ABS08498 Human genome-deriv
c1000 16 0.5 471 21 AAC44472 Zea mays DNA fragm
```

## ALIGNMENTS

```
RESULT 1
AAAI5299 standard; DNA: 3204 BP.
```

```
XX AAAI5299;
```

```
XX 04-SEP-2000 (first entry)
```

```
XX DNA encoding a polypeptide of a Neisseria pathogenic strain.
```

```
XX Pathogenic strain; Neisseria; vaccine; Neisseria infection; ss.
```

```
XX Neisseria meningitidis.
```

```
XX Key Location/Qualifiers
FT CDS 1..3204
FT /*tag= a
```

```
XX WO200026375-A2.
```

```
XX 11-MAY-2000.
```

```
XX 28-OCT-1999; 99WO-FR02643.
```

```
XX 30-OCT-1998; 98FR-0013693.
```

```
XX
```

```
PA (INMR ) PASTEUR MERIEUX SERUMS & VACCINS SA.
PA (INMR ) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI Aujame L, Bouchardon A, Renauld-Mongenie G, Rokbi B, Nassif X;
PI Tinsley C, Perrin A;
XX
DR WPI; 2000-365622/31.
DR P-PSDB; AAY93268.
XX
PT New polypeptide specific for pathogenic Neisseria useful in therapeutic
PT or preventative vaccines and for diagnosis -
XX
XX Claim 2; Page 61-65; 187pp; French.
XX
CC The present sequence encodes a protein that is specific for pathogenic
CC strains of Neisseria. The polynucleotides, polypeptides, or their
CC antigenic fragments, are used in vaccines to treat or protect against
CC Neisseria infections, particularly by N. meningitidis. The
CC polynucleotide sequence is also used for recombinant production of
CC the polypeptide and to produce attenuated Neisseria strains that
CC overexpress it, or express it in a non-toxic mutant form.
XX
SQ Sequence 3204 BP: 831 A; 909 C; 899 G; 565 T; 0 other;
Query Match 100.0%; Score 3204; DB 21; Length 3204;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCCAAGACCCCAACCTTCCTACAAAACCTTTCAAAACCGGTCGCATGCGGTAGCT 60
Db 1 ATGCCAAGACCCCAACCTTCCTACAAAACCTTTCAAAACCGGTCGCATGCGGTAGCT 60
QY 61 GTTGCAACAACACTTTCGCTCTAGCGCGCGGCGGCACTTCTGCGCCGACTTC 120
Db 61 GTTGCAACAACACTTTCGCTCTAGCGCGCGGCGGCACTTCTGCGCCGACTTC 120
QY 121 AATGACGCGGCACCGGTATCGGACCAACAGCAGCAACAAACGCGAAATTCAGCAGCA 180
Db 121 AATGACGCGGCACCGGTATCGGACCAACAGCAGCAACAAACGCGAAATTCAGCAGCA 180
QY 181 GTATCTTACGCGCGGTATCAAGAACGAATGTGCAAGACAGCAAGCATGCTGTGCCGGT 240
Db 181 GTATCTTACGCGCGGTATCAAGAACGAATGTGCAAGACAGCAAGCATGCTGTGCCGGT 240
QY 241 CGGGATGAGCTTGGGTTACAGACAGGATGCCAAATCAATGCCGCCGCCGCCGAATCTG 300
Db 241 CGGGATGAGCTTGGGTTACAGACAGGATGCCAAATCAATGCCGCCGCCGCCGAATCTG 300
QY 301 CATACCGGAGACTTTTACAAACCCAAATGACGCATACAGAATTTGATCAACCTCAAACT 360
Db 301 CATACCGGAGACTTTTACAAACCCAAATGACGCATACAGAATTTGATCAACCTCAAACT 360
QY 361 GCAATTGAAGCAGGCTATACAGGACGCGGGTAGAGGTAGGTATCGTGATACAGCGGNA 420
Db 361 GCAATTGAAGCAGGCTATACAGGACGCGGGTAGAGGTAGGTATCGTGATACAGCGGNA 420
QY 421 TCCGTCGCGCAGCATATCCTTTCCGAACTGTATGCGAGAAAAGAACACGGCTATACGAA 480
Db 421 TCCGTCGCGCAGCATATCCTTTCCGAACTGTATGCGAGAAAAGAACACGGCTATACGAA 480
QY 481 AATTACAAAACACTATACGCGGTATATGCGGAAGAGCGCCTGAAGACGAGGCGGTA 540
Db 481 AATTACAAAACACTATACGCGGTATATGCGGAAGAGCGCCTGAAGACGAGGCGGTA 540
QY 541 GACATTAAAGCTTCTTCGACCATGAGCCGTTATAGAGACTGAAGCAAGCCGACGGAT 600
Db 541 GACATTAAAGCTTCTTCGACCATGAGCCGTTATAGAGACTGAAGCAAGCCGACGGAT 600
QY 601 ATCCGCCACGTAAGAAATCGACACATCGATGTGGTCTCCCATATATTTCGCGGCGT 660
Db 601 ATCCGCCACGTAAGAAATCGACACATCGATGTGGTCTCCCATATATTTCGCGGCGT 660
QY 661 TCCGTGGACGCGCAGACCTGTCAGGCGGTATTGCGCCGATGCGACGCTACACATAATGAAT 720
```



[illegible]

Db	1741	AAAGCGATCTGCAGCTGGCGGCGGAAGGTACGCTGTACACACGTTTGGGCAAACTGCTG	1800
Qy	1801	AAAGTGGACGGTACGGCGATGACCGGCGCAAGCTGTACATGTCCGCACAGCGGCAAAAGG	1860
Db	1801	AAAGTGGACGGTACGGCGATGACCGGCGCAAGCTGTACATGTCCGCACAGCGGCAAAAGG	1860
Qy	1861	CGAGGCTATCTCAACCGTACCGGACAACTGTTCCCTTCCCTGAGTCCGCCCAAAATCGGG	1920
Db	1861	CGAGGCTATCTCAACCGTACCGGACAACTGTTCCCTTCCCTGAGTCCGCCCAAAATCGGG	1920
Qy	1921	CGGGATTATCTTCTTCACAAACATCGAAACCGACGGTGGTCTGCTGGCTTCCCTCGAC	1980
Db	1921	CGGGATTATCTTCTTTCACAAACATCGAAACCGACGGTGGTCTGCTGGCTTCCCTCGAC	1980
Qy	1981	AGCGTCTGAAAAAAGCGGGCGACGTGAAGCGACACGCTCTCCCTATTATGTCTCGTCGGCGC	2040
Db	1981	AGCGTCTGAAAAAAGCGGGCGACGTGAAGCGACACGCTCTCCCTATTATGTCTCGTCGGCGC	2040
Qy	2041	AATCGGCGACGGACTGTTTCGGCAGCGGACATTTCCGCGCCGCCGGTCTGAAACACGCC	2100
Db	2041	AATCGGCGACGGACTGTTTCGGCAGCGGACATTTCCGCGCCGCCGGTCTGAAACACGCC	2100
Qy	2101	GTAGAACAGGGCGGCACAACTCTGAAAAACCTGATGTCGAACTGATGGTTCGAACCTCGATCGAATCA	2160
Db	2101	GTAGAACAGGGCGGCACAACTCTGAAAAACCTGATGTCGAACTGATGGTTCGAACCTCGAATCA	2160
Qy	2161	TCCGCAACACCCGAGACGGTTGAAACTCGGGCGCGCACGACACAGATATCGCGGGCATC	2220
Db	2161	TCCGCAACACCCGAGACGGTTGAAACTCGGGCGCGCACGACACAGATATCGCGGGCATC	2220
Qy	2221	CGCCCTTACGGCGCAACTTTTCCGCGCAGCGGACGCGTACAGCATGCGAATGCCCGCGAC	2280
Db	2221	CGCCCTTACGGCGCAACTTTTCCGCGCAGCGGACGCGTACAGCATGCGAATGCCCGCGAC	2280
Qy	2281	GGTGTAGCATCTTCAACAGTCTCGCGCTACCGTCTATGCGGACAGTACCGCGCGCCAT	2340
Db	2281	GGTGTAGCATCTTCAACAGTCTCGCGCTACCGTCTATGCGGACAGTACCGCGCGCCAT	2340
Qy	2341	GCCGATATGCAGGACCGCGGCTGAAAGCCGTATCGAGCGGTTGGACACAAACGCTACG	2400
Db	2341	GCCGATATGCAGGACCGCGGCTGAAAGCCGTATCGAGCGGTTGGACACAAACGCTACG	2400
Qy	2401	GGTCTGCGGCTATCGGCGAAACCCCAACAGGACGGTGGAACTGGGAAACAGGGCGGTGT	2460
Db	2401	GGTCTGCGGCTATCGGCGAAACCCCAACAGGACGGTGGAACTGGGAAACAGGGCGGTGT	2460
Qy	2461	GAAAGCAAAATGCGCGCAGTACCCAAACCGTTCGGCTGGCTTCCCGCAAAACCGCGGAAAT	2520
Db	2461	GAAAGCAAAATGCGCGCGAGTACCCAAACCGTTCGGCTGGCTTCCCGCAAAACCGCGGAAAT	2520
Qy	2521	ACGACAGCGCGCACACTGGGACATGGGACACAGCACATGGAGCGGAAACAGTGCAAAT	2580
Db	2521	ACGACAGCGCGCACACTGGGACATGGGACACAGCACATGGAGCGGAAACAGTGCAAAT	2580
Qy	2581	GCAAAAACCGACAGCATTAGTCTGTTTGACGACATACGGCAGCATGCGGGCGATATCGGC	2640
Db	2581	GCAAAAACCGACAGCATTAGTCTGTTTGACGACATACGGCAGCATGCGGGCGATATCGGC	2640
Qy	2641	TATCTCAAAGCCCTGTTCTCCTTACGACAGCTACAAAACAGCATCAGCGCAGCACCGGT	2700
Db	2641	TATCTCAAAGCCCTGTTCTCCTTACGACAGCTACAAAACAGCATCAGCGCAGCACCGGT	2700
Qy	2701	CGGACCAACATCGGAAGCAGCGTCAAGGCACGCTGATCCAGCTGGGGCGACTGGGC	2760
Db	2701	CGGACCAACATCGGAAGCAGCGTCAAGGCACGCTGATCCAGCTGGGGCGACTGGGC	2760
Qy	2761	GGTGTCAACGTTCCGTTTTCGCGCAACGGGAGATTGTGACGGTCTCAAGCGGCTCTCGCTAC	2820
Db	2761	GGTGTCAACGTTCCGTTTTCGCGCAACGGGAGATTGTGACGGTCTCAAGCGGCTCTCGCTAC	2820
Qy	2821	GACCTGTCTAAACAGGATGATTCGCCGCAAAAGGACGTGCTTTGGGCTGGAGCGGCAAC	2880
Db	2821	GACCTGTCTAAACAGGATGATTCGCCGCAAAAGGACGTGCTTTGGGCTGGAGCGGCAAC	2880

QY 2881 AGCCTCACTGAAGGACACACTGGTTCGGACTCGCGGGTCTGAAGTGTGCGAACCCCTTGAGC 2940  
|||||  
Db 2881 AGCCTCACTGAAGGACACACTGGTTCGGACTCGCGGGTCTGAAGTGTGCGAACCCCTTGAGC 2940  
QY 2941 GATAAAGCGTCTCTGTTTCAACAGCGCGGCGTGGAAACGCGACTGAACGAGCGGCGACTAC 3000  
Db 2941 GATAAAGCGTCTCTGTTTCAACAGCGCGGCGTGGAAACGCGACTGAACGAGCGGCGACTAC 3000  
QY 3001 ACGGTAACGGCGGCGCTTTTACCGCGCGACTGAGCAACCGCGGAAGACGGGGCGACGCAAT 3060  
Db 3001 ACGGTAACGGCGGCGCTTTTACCGCGCGACTGAGCAACCGCGGAAGACGGGGCGACGCAAT 3060  
QY 3061 ATGCCGCACACCGCGCTGCTGCGGTCTGCGGCGGATGTGGAATTCGCGAACGGCTGG 3120  
Db 3061 ATGCCGCACACCGCGCTGCTGCGGTCTGCGGCGGATGTGGAATTCGCGAACGGCTGG 3120  
QY 3121 AACGGCTTGGCAGCTTACAGCTACGCGCGTTCGAACACAGTACGGCAACACAGCGGACGA 3180  
Db 3121 AACGGCTTGGCAGCTTACAGCTACGCGCGTTCGAACACAGTACGGCAACACAGCGGACGA 3180  
QY 3181 CTCGGCGTAGGCTACCGGTTCTGA 3204  
Db 3181 CTCGGCGTAGGCTACCGGTTCTGA 3204

RESULT 2

AAZ53301  
ID AAZ53301 standard; DNA; 1365 BP.  
XX  
AC AAZ53301;  
XX  
DT 21-MAR-2000 (first entry)  
DE Neisseria meningitidis ORF 140 partial DNA sequence SEQ ID NO:551.  
XX  
KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;  
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;  
KW antibacterial; gene therapy; ds.  
XX  
OS Neisseria meningitidis.  
PN WO9957280-A2.  
PD 11-NOV-1999.  
PF 30-APR-1999; 99WO-US09346.  
PR 01-MAY-1998; 98US-0083758.  
PR 31-JUL-1998; 98US-0094869.  
PR 02-SEP-1998; 98US-0098994.  
PR 02-SEP-1998; 98US-0099062.  
PR 09-OCT-1998; 98US-0103749.  
PR 09-OCT-1998; 98US-0103794.  
PR 09-OCT-1998; 98US-0103796.  
PR 25-FEB-1999; 99US-0121528.  
XX  
PA (CHIR ) CHIRON CORP.  
PA (GENO-) INST GENOMIC RES.  
XX  
PI Fraser C, Galeotti C, Grandi G, Hickey E, Massignani V, Mora M;  
PI Petersen J, Pizza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;  
PI Tettelin H, Venter JC;  
XX  
XX WPI; 2000-062150/05.  
DR P-PSDB; AAY74539.  
XX  
PT Novel Neisserial polypeptides predicted to be useful antigens for  
PT vaccines and diagnostics  
XX  
PS Claim 7; Page 396-397; 1453pp; English.  
XX  
QC AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941

CC represent novel *Neisseria meningitidis* and *N. gonorrhoeae* polynucleotides  
CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ5473 represent  
CC PCR primers used in the exemplification of the present invention. The  
CC polypeptides, the polynucleotides, antibodies and compositions of  
CC the invention can be used as vaccines, as diagnostic reagents, and as  
CC immunogenic compositions. The polypeptides can be used in the  
CC manufacture of medicaments for treating or preventing infection due to  
CC *Neisseria meningitidis* (e.g. meningitis and septicaemia), to detect the  
CC presence of *Neisseria meningitidis*, or to raise antibodies. They may also  
CC be used to screen for agonists or antagonists, which may themselves  
CC have use as antibacterial agents. The polynucleotides of the invention  
CC may also be used in gene therapy protocols.  
XX  
SQ Sequence 1365 BP; 319 A; 410 C; 418 G; 218 T; 0 other;  
Query Match 30.9%; Score 990; DB 21; Length 1365;  
Best Local Similarity 99.5%; Pred. No. 0;  
Matches 1340; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
QY 1858 GGGGCGAGGTATCTCAACCGTACCGGACAACTGTTCCCTTCTGAGTGCCGCCAAAATC 1917  
Db 19 GGGGCGAGGTATCTCAACCGTACCGGACAACTGTTCCCTTCTGAGTGCCGCCAAAATC 78  
QY 1918 GGGGCGAGGTATCTCAACCGTACCGGACAACTGTTCCCTTCTGAGTGCCGCCAAAATC 1977  
Db 79 GGGGCGAGGTATCTCAACCGTACCGGACAACTGTTCCCTTCTGAGTGCCGCCAAAATC 138  
QY 1978 GACAGGTCGAAAAACAGCGGAGTGAAGCGGACAGCTGCTCTATATATCTGCGTGC 2037  
Db 139 GACAGGTCGAAAAACAGCGGAGTGAAGCGGACAGCTGCTCTATATATCTGCGTGC 198  
QY 2038 GGCAATGCGGCGACGACTGCTTCGGCGAGCGGACATTCGCGCGCGCGGTCTGAAACAC 2097  
Db 199 GGCAATGCGGCGACGACTGCTTCGGCGAGCGGACATTCGCGCGCGCGGTCTGAAACAC 258  
QY 2098 GCGGTAGAACAGGCGGCGCAATCTGAAAAACCTGATGTGCGAACTGGATGCTCCGAA 2157  
Db 259 GCGGTAGAACAGGCGGCGCAATCTGAAAAACCTGATGTGCGAACTGGATGCTCCGAA 318  
QY 2158 TCATCCGCAACACCGGAGACGGTTGAAACTGCGGCGCGGCGGACGACAGATATGCGGCG 2217  
Db 319 TCATCCGCAACACCGGAGACGGTTGAAACTGCGGCGCGGCGGACGACAGATATGCGGCG 378  
QY 2218 ATCCGCGCGCTACGCGCAACTTTCGCGCGAGCGGACGCTACAGCATGCGGCGCG 2277  
Db 379 ATCCGCGCGCTACGCGCAACTTTCGCGCGAGCGGACGCTACAGCATGCGGCGCG 438  
QY 2278 GACGTGTACGATCTTCAACAGTCTCGCGGCTACCGTCTATGCGGACAGTACCGCGCG 2337  
Db 439 GACGTGTACGATCTTCAACAGTCTCGCGGCTACCGTCTATGCGGACAGTACCGCGCG 498  
QY 2338 CATGCGGATATGCGGCGGCGGCTGAAAGCGGCTATGCGGCGGCTGCGGCGGCGGAA 2397  
Db 499 CATGCGGATATGCGGCGGCGGCTGAAAGCGGCTATGCGGCGGCTGCGGCGGCGGAA 558  
QY 2398 ACGGTCTGCGGCTCATCGCGCAACCCAAACAGGACGGTGGAAACGTTGGAAACAGGCGGT 2457  
Db 559 ACGGTCTGCGGCTCATCGCGCAACCCAAACAGGACGGTGGAAACGTTGGAAACAGGCGGT 618  
QY 2458 GTTGAAGGCAAAATGCGGCGGCGGCTACCCAAACCGTTCGCGGCGGCGGCGGCGGAA 2517  
Db 619 GTTGAAGGCAAAATGCGGCGGCGGCTACCCAAACCGTTCGCGGCGGCGGCGGCGGAA 678  
QY 2518 AATAGCAGCAGCGCGGCGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2577  
Db 679 AATAGCAGCAGCGCGGCGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 738  
QY 2578 AATGCAAAACCGCAGCAGTATGCTGTTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2637  
Db 739 AATGCAAAACCGCAGCAGTATGCTGTTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 798  
QY 2638 GCGTATCTCAAGGCGGCTGTTCTTCCAGGCGGCTCAAAAACAGCATACGCGCGGCGGCGG 2697  
|||||

Db 799 GGCTATCTCAAAGCCCTGTTCTCTACGGAGCTGTACAAAACAGCATCATGCCCGCAGCAC 858  
QY 2698 GGTGCGGACGAACATGCGAAGGACAGCTCAACGGCAGCGTGTATGACAGCTGGGCGCACTG 2757  
Db 859 GGTGCGGACGAACATGCGAAGGACAGCTCAACGGCAGCGTGTATGACAGCTGGGCGCACTG 918  
QY 2758 GGGCGGTGTCAACGTTCCGTTTGGCGCAACGGGAGATTGTGACGGTTCGAAGCGGTTCTCGC 2817  
Db 919 GGGCGGTGTCAACGTTCCGTTTGGCGCAACGGGAGATTGTGACGGTTCGAAGCGGTTCTCGC 978  
QY 2818 TAGCACTGTCTAAACAGGATGATTCGCCGAAAAAGGAGTCTTTGGGCTGGAGCGG 2877  
Db 979 TAGCACTGTCTAAACAGGATGATTCGCCGAAAAAGGAGTCTTTGGGCTGGAGCGG 1038  
QY 2878 AACAGCTCTACTGAAGCAGACACTGGTCGGACTCGCGGCTCTGAAGCTGTGCAACCTTG 2937  
Db 1039 AACAGCATCACTGAAGCAGACACTGGTCGGACTCGCGGCTCTGAAGCTGTGCAACCTTG 1098  
QY 2938 AGCGATAAAGCCGTCCTGTTTGAACGGCGGGGCTGGAACCGCAGCTGAACGGACGCGAC 2997  
Db 1099 AGCGATAAAGCCGTCCTGTTTGAACGGCGGGGCTGGAACCGCAGCTGAACGGACGCGAC 1158  
QY 2998 TACAGGTACGGCGGCTTTACGGCGGCGACTGTCAGCAACCGGCAAGACGGGGCAGC 3057  
Db 1159 TACAGGTACGGCGGCTTTACGGCGGCGACTGTCAGCAACCGGCAAGACGGGGCAGC 1218  
QY 3058 AATATCGCGCACACCCGCTGTTGCCGCTTGGCGCGGATGTCGAATTCGGCAACGGC 3117  
Db 1219 AATATCGCGCACACCCGCTGTTGCCGCTTGGCGCGGATGTCGAATTCGGCAACGGC 1278  
QY 3118 TGAACGGCTTGGCAGCTTACAGCTACGGCGGTTTCAAAACAGTACGGCAACACACAGCGGA 3177  
Db 1279 TGAACGGCTTGGCAGCTTACAGCTACGGCGGTTTCAAAACAGTACGGCAACACACAGCGGA 1338  
QY 3178 CGAGTGGCGTAGCTACCGGTTCTGA 3204  
Db 1339 CGAGTGGCGTAGCTACCGGTTCTGA 1365

RESULT 3  
AAZ53300  
ID AAZ53300 standard; DNA; 1365 BP.  
XX AC AAZ53300;  
XX XX  
XX 21-MAR-2000 (first entry)  
XX DE Neisseria meningitidis ORF 140 partial DNA sequence SEQ ID NO:549.  
XX KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;  
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;  
KW antibacterial; gene therapy; ds.  
XX OS Neisseria meningitidis.  
XX PN W09957280-A2.  
XX XX  
XX PD 11-NOV-1999.  
XX XX  
XX PF 30-APR-1999; 99WO-0509346.  
XX XX  
XX PR 01-MAY-1998; 98US-0083758.  
XX PR 31-JUL-1998; 98US-0094869.  
XX PR 02-SEP-1998; 98US-0098994.  
XX PR 02-SEP-1998; 98US-0099062.  
XX PR 09-OCT-1998; 98US-0103749.  
XX PR 09-OCT-1998; 98US-0103796.  
XX PR 25-FEB-1999; 99US-0121528.  
XX XX  
XX PA (CHIR ) CHIRON CORP.  
XX PA (GENO-) INST GENOMIC RES.  
XX XX

PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;  
PI Petersen J, Pizza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;  
PI Tettelin H, Venter JC;  
XX WPI; 2000-062150/05.  
DR P-PSDB; AAY74538.  
XX Novel Neisserial polypeptides predicted to be useful antigens for  
PT vaccines and diagnostics  
PT  
XX Claim 7: Page 395; 1453pp; English.  
XX  
CC AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941  
CC represent novel Neisseria meningitis and N. gonorrhoeae polynucleotides  
CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ5473 represent  
CC PCR primers used in the exemplification of the present inventions. The  
CC polypeptides, the polynucleotides, antibodies and compositions of  
CC the invention can be used as vaccines, as diagnostic reagents, and as  
CC immunogenic compositions. The polypeptides can be used in the  
CC manufacture of medicaments for treating or preventing infection due to  
CC Neisserial bacteria (e.g. meningitis and septicaemia), to detect the  
CC presence of Neisseria bacteria, or to raise antibodies. They may also  
CC be used to screen for agonists or antagonists, which may themselves  
CC have use as antibacterial agents. The polynucleotides of the invention  
CC may also be used in gene therapy protocols.  
XX  
SQ Sequence 1365 BP; 317 A; 413 C; 421 G; 214 T; 0 other;  
Query Match 26.0%; Score 834; DB 21; Length 1365;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 1234; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 1963 CTGCTGGCTTCCCTGACAGCGTGCAGAAAAACAGCGGAGTGAAGCGACAGCGTGTCC 2022  
Db 124 CTGCTGGCTTCCCTGACAGCGTGCAGAAAAACAGCGGAGTGAAGCGACAGCGTGTCC 183  
QY 2023 TATTATGTCTCGTCGGCAATGCGGCAGCGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2082  
Db 184 TATTATGTCTCGTCGGCAATGCGGCAGCGACTGCTTCGGCAGCGGCACATTCCGCGCCC 243  
QY 2083 GCCGGTCTGAACACACGCCGTGAGAACAGGGCGGAGCAATCTGGAACACCTGTATGGTGA 2142  
Db 244 GCCGGTCTGAACACACGCCGTGAGAACAGGGCGGAGCAATCTGGAACACCTGTATGGTGA 303  
QY 2143 CTGATGCTTCCGAATCATCCGCAACACCGAGAGCGTTGAAACTCGGCGCGGACCGC 2202  
Db 304 CTGATGCTTCCGAATCATCCGCAACACCGAGAGCGTTGAAACTCGGCGCGGACCGC 363  
QY 2203 ACAGATATGCGGGCATCCGCCCTACGGCGCAACTTTCGGCGCAGCGGCGGTACAG 2262  
Db 364 ACAGATATGCGGGCATCCGCCCTACGGCGCAACTTTCGGCGCAGCGGCGGTACAG 423  
QY 2263 CATGCGAATGCGCGGACGGTGTACGCATCTTCAACAGTCTCGCGCTACCGCTATGCC 2322  
Db 424 CATGCGAATGCGCGGACGGTGTACGCATCTTCAACAGTCTCGCGCTACCGCTATGCC 483  
QY 2323 GACAGTACCGCGCCCATGCCGATATGCGGAGCGCGGTGAAGCCGTATCGGACGG 2382  
Db 484 GACAGTACCGCGCCCATGCCGATATGCGGAGCGCGGTGAAGCCGTATCGGACGG 543  
QY 2383 TTGGACCAACAGCTACGGGTCTGCGGTCTATCGCGCAACCCCAACAGGAGCGTGAACG 2442  
Db 544 TTGGACCAACAGCGGCGGTCTGCGGTCTATCGCGCAACCCCAACAGGAGCGTGAACG 603  
QY 2443 TGGGAACAGGGCGGTGTGAAGGCAAAATGCGGGCAGTAGTACCCAAACCGTCGGCATTTGCC 2502  
Db 604 TGGGAACAGGGCGGTGTGAAGGCAAAATGCGGGCAGTAGTACCCAAACCGTCGGCATTTGCC 663  
QY 2503 GCGAAAACCGGCGAAATATGACAGACGAGCGCGGCACACTGGGGCATGGGACACACATGG 2562  
Db 664 GCGAAAACCGGCGAAATATGACAGACGAGCGCGGCACACTGGGGCATGGGACACACATGG 723  
QY 2563 AGCGAAAACAGTGCAAAATGCAAAACCGAGCATTAGTCTGTTTGCAGGCATACGGCAC 2622

Db 724 AGCGAAACAGTGCATAATCAAAACCCGACAGCATTAGTCTGTTTCAGCGCATACGGCAC 783  
Qy 2623 GATGCGGGCGATATCGCTATCTCAAGAGCCTCTTCTCTACGAGCGCTACAAAACAGC 2682  
Db 784 GATGCGGGCGATATCGCTATCTCAAGAGCCTCTTCTCTACGAGCGCTACAAAACAGC 843  
Qy 2683 ATCAGCCGCGACCGGTGCGGACGACATGCGGAAGGAGCGGTCAACGGCACGCTGATG 2742  
Db 844 ATCAGCCGCGACCGGTGCGGACGACATGCGGAAGGAGCGGTCAACGGCACGCTGATG 903  
Qy 2743 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGGCGCAACGGGAGTTTGACGGTC 2802  
Db 904 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGGCGCAACGGGAGTTTGACGGTC 963  
Qy 2803 GAAGGCGGTCTCGCTTACGACCTGCTCAACAGAGGATGCTTCGCCGAAAGAGCGAGTGTCT 2862  
Db 964 GAAGGCGGTCTCGCTTACGACCTGCTCAACAGAGGATGCTTCGCCGAAAGAGCGAGTGTCT 1023  
Qy 2863 TTGGGCTGAGGCGGCAACACGCTCACTGAAGGCACACTGGTGGACTCGCGGTCTGAAG 2922  
Db 1024 TTGGGCTGAGGCGGCAACACGCTCACTGAAGGCACGCTGTGCGACTCGCGGTCTGAAG 1083  
Qy 2923 CTGTCCACACCTTGAGCGGATGAAGCCGTCCTGTTTGAAGCGGCGGTGGAACCGCAC 2982  
Db 1084 CTGTCCACACCTTGAGCGGATGAAGCCGTCCTGTTTGAAGCGGCGGTGGAACCGCAC 1143  
Qy 2983 CTGAAGGAGCGGACTTACACGCTTAACGGCGGCTTTACCGCGCGACTCAACGACCGGC 3042  
Db 1144 CTGAAGGAGCGGACTTACACGCTTAACGGCGGCTTTACCGCGCGACTCAACGACCGGC 1203  
Qy 3043 AAGACGGGGGACGCAATATGCCGCACACCCGCGTGGTTGCGGGCTGGGCGCGGATGTC 3102  
Db 1204 AAGACGGGGGACGCAATATGCCGCACACCCGCTGTGGTTGCGGGCTGGGCGCGGATGTC 1263  
Qy 3103 GAATTCGGAACGGCTGGAAGCGCTTGGACAGGTTACAGCTACGCGGTTCCAAACAGTAC 3162  
Db 1264 GAATTCGGAACGGCTGGAAGCGCTTGGACAGGTTACAGCTACGCGGTTCCAAACAGTAC 1323  
Qy 3163 GGCACACACGCGGAGGAGTTCGCGGTAGGCTACCGGTTCTGA 3204  
Db 1324 GGCACACACGCGGAGGAGTTCGCGGTAGGCTACCGGTTCTGA 1365

RESULT 4

AA543905  
ID AA543905 standard; DNA; 4218 BP.  
XX AA543905;  
AC AA543905;  
XX  
DT 18-DEC-2001 (first entry)  
XX  
DE Neisseria meningitidis fusion protein 961cL-983 DNA.  
XX

KW Neisseria gonorrhoeae; leader peptide; fusion protein; ORF46.1; ds;  
KW Neisserial protein.  
XX

OS Neisseria meningitidis.  
OS Synthetic.

XX WO200164922-A2.  
XX

XX 07-SEP-2001.  
XX

XX 28-FEB-2001; 2001WO-1B00452.  
XX

XX 28-FEB-2000; 2000GB-0004695.  
PR 13-NOV-2000; 2000GB-0027675.

XX (CHIR-) CHIRON SPA.  
XX

XX Arico MB, Comanducci M, Galeotti C, Masignani V, Guilianini MM;  
PI Pizza M;  
QI

XX WPI; 2001-582163/65.  
DR P-PSDB; AAU27608.  
XX  
PT Producing heterologous proteins from Neisseria meningitidis and N.  
PT gonorrhoeae -  
XX  
PS Example 23; Page 75-76; 119pp; English.  
XX  
CC The invention relates to methods for the heterologous expression of  
CC Neisserial proteins from Neisseria meningitidis and Neisseria  
CC gonorrhoeae. At least one domain in the protein is deleted, e.g. the  
CC leader peptide, and may be replaced by a domain from a different protein  
CC to make a fusion protein, in order to enhance heterologous expression of  
CC Neisserial proteins. Also, a region of a protein, such as a poly-glycine  
CC stretch, can be mutated to enhance expression. The proteins used in the  
CC processes include ORF46.1, 287, 741, 919, 953, 961 and 983. Sequences  
CC AA543868-AA543905 represent DNA molecules encoding Neisserial proteins  
CC and peptide regions of proteins of the invention.  
XX  
SQ Sequence 4218 BP; 1196 A; 1149 C; 1130 G; 743 T; 0 other;

Query Match 26.0%; Score 834; DB 22; Length 4218;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 1234; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1963 CTGCTGGCTTCCTCGACAGCGTCGAAAAACAGCGGCGAGTGAAGGCGACACGCTGTCC 2022  
Db 2971 CTGCTGGCTTCCTCGACAGCGTCGAAAAACAGCGGCGAGTGAAGGCGACACGCTGTCC 3030  
Qy 2023 TATTATGTCGCTCGCGGCAATGCGGCACGGACTGCTGGGACGGGCACATTCCTCCGCGCC 2082  
Db 3031 TATTATGTCGCTCGCGGCAATGCGGCACGGACTGCTGGGACGGGCACATTCCTCCGCGCC 3090  
Qy 2083 GCCGCTCTGAACACCGCGTAGAACAGGCGGCGACCAATCTGGAAAACTGATGTCGAA 2142  
Db 3091 GCCGCTCTGAACACCGCGTAGAACAGGCGGCGACCAATCTGGAAAACTGATGTCGAA 3150  
Qy 2143 CTGGATGCTCCGGAATCATCCGCAACACCGAGAGCGGTTGAACTCGCGCCCGCACCGC 2202  
Db 3151 CTGGATGCTCCGGAATCATCCGCAACACCGAGACGGTTGAACTCGCGGCGCACCGC 3210  
Qy 2203 ACAGATATGCGGGGATCGCCCTTACGGCGCAACTTTCCGCGCAGCGGACGCCGTATAC 2262  
Db 3211 ACAGATATGCGGGGATCGCCCTTACGGCGCAACTTTCCGCGCAGCGGACGCCGTATAC 3270  
Qy 2263 CATGCGAATGCGCGCGAGGTGTACGCATCTTCAACAGTCTCGCGCTACCGCTATGCC 2322  
Db 3271 CATGCGAATGCGCGCGAGGTGTACGCATCTTCAACAGTCTCGCGCTACCGCTATGCC 3330  
Qy 2323 GACGATACGCGCGCCATGCCGATATGAGGAGCGCGGCTGAAAGCGGTATCGGACGG 2382  
Db 3331 GACGATACGCGCGCCATGCCGATATGAGGAGCGCGGCTGAAAGCGGTATCGGACGG 3390  
Qy 2383 TTGGACCAACACGCTACGGGTCTGCGCGTCAATCGCGCAAAACCAACAGGACGGTGGAAAG 2442  
Db 3391 TTGGACCAACACGCGGCTGCGCGTCAATCGCGCAAAACCAACAGGACGGTGGAAAG 3450  
Qy 2443 TGGGAACAGGCGGTGTTGAAGGCAAAATGCGCGCAGTACCCAAACCGCTCGGCATTGCC 2502  
Db 3451 TGGGAACAGGCGGTGTTGAAGGCAAAATGCGCGCAGTACCCAAACCGCTCGGCATTGCC 3510  
Qy 2503 GCGAAACCGGCGAAATACGACAGCGCGCACACTGGGCGATGGGACACACATGG 2562  
Db 3511 GCGAAACCGGCGAAATACGACAGCGCGCACACTGGGCGATGGGACACACATGG 3570  
Qy 2563 AGCGAAAAAGTGCATAATGCAAAAAACCGACAGCATTAGTCTGTTTTCAGGCGATACGGCAC 2622  
Db 3571 AGCGAAAAAGTGCATAATGCAAAAAACCGACAGCATTAGTCTGTTTTCAGGCGATACGGCAC 3630  
Qy 2623 GATGCGGGCGATATCGGCTATCTCAAGAGCCCTGTTCTCTACGAGCGCTACAAAAACAGC 2682  
Db 3631 GATGCGGGCGATATCGGCTATCTCAAGAGCCCTGTTCTCTCTACGAGCGCTACAAAAACAGC 3690





Db	8342	GGCAAAACCGCGAATAATAGCAGCAGCGCGGCACACTGGGCAATGGAGCGCAGCATGG	8283
QY	2563	AGCGAAACAGTGCAATGCAAAACCGCAGCAGCATTTAGTCTGTTTTCAGGCATACGGCAC	2622
Db	8282	AGCGAAACAGTGCAATGCAAAACCGCAGCAGCATTTAGTCTGTTTTCAGGCATACGGCAC	8223
QY	2623	GATGGGGCGATATGCGCTATCTCAAAGCCCTGTTCTCTACGGAGCTACAAAAACAGC	2682
Db	8222	GATGGGGCGATATGCGCTATCTCAAAGCCCTGTTCTCTACGGAGCTACAAAAACAGC	8163
QY	2683	ATCAGCCGACGACCGCTCGGACGCAACATGCGGAGGCGAGCTCAACGCGCAGCTGATG	2742
Db	8162	ATCAGCCGACGACCGCTCGGACGCAACATGCGGAGGCGAGCTCAACGCGCAGCTGATG	8103
QY	2743	CAGCTGGGCGCACTGGGCGGTGTCAACGTTTCGTCGCAACGGGAGATTTCAGCGTC	2802
Db	8102	CAGCTGGGCGCACTGGGCGGTGTCAACGTTTCGTCGCAACGGGAGATTTCAGCGTC	8043
QY	2803	GAAGGGGTCTGCGTACGACCTGTCTAAACAGGATGCAATTCGCCGAAAAAGGCAAGTGT	2862
Db	8042	GAAGGGGTCTGCGTACGACCTGTCTAAACAGGATGCAATTCGCCGAAAAAGGCAAGTGT	7983
QY	2863	TTGGGCTGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGTCTGAAG	2922
Db	7982	TTGGGCTGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGTCTGAAG	7923
QY	2923	CTGTGCGCAACCCCTTGAGCGATAAAGCCGTCCTGTTTGCACACGGCGGCGTGAACGCGAC	2982
Db	7922	CTGTGCGCAACCCCTTGAGCGATAAAGCCGTCCTGTTTGCACACGGCGGCGTGAACGCGAC	7863
QY	2983	CTGAACGGACGGCACTACACGGTAACGGCGGCTTTTACCGGCGCACTCAGCAACCGCG	3042
Db	7862	CTGAACGGACGGCACTACACGGTAACGGCGGCTTTTACCGGCGCACTCAGCAACCGCG	7803
QY	3043	AGACGGGGGCGCAATATGCGCACACCCCGCTGTTGCCGGTCTGGCGCGGATGTC	3102
Db	7802	AGACGGGGGCGCAATATGCGCACACCCCGCTGTTGCCGGTCTGGCGCGGATGTC	7743
QY	3103	GAATTCGGCAACGGCTGGAAACGGCTTGGCACCTTACAGCTACGCGGTTTCCAAACAGTAC	3162
Db	7742	GAATTCGGCAACGGCTGGAAACGGCTTGGCACCTTACAGCTACGCGGTTTCCAAACAGTAC	7683
QY	3163	GGCAACACAGCGGAGCTGGCGGTAGGCTACCGGTTCTGA	3204
Db	7682	GGCAACACAGCGGAGCTGGCGGTAGGCTACCGGTTCTGA	7641

RESULT 7

ID	AAF21612	AAF21612 standard; DNA; 349980 BP.
XX		
AC	AAF21612;	
XX		
DT	13-MAR-2001	(first entry)
XX		
DE		Neisseria meningitidis B nucleotide sequence SEQ ID NO:113.
XX		
KW		Neisseria meningitidis; Neisseria gonorrhoeae; immunogenic; vaccine;
KW		diagnosis; antigen; detection; infection; gene therapy; antibacterial;
KW		ds.
XX		
OS		Neisseria meningitidis.
XX		
PN	WO200066791-A1.	
XX		
PD	09-NOV-2000.	
XX		
PF	08-MAR-2000; 2000WO-US05928.	
XX		
PR	30-APR-1999; 98US-0132068.	
PR	08-OCT-1999; 99WO-US23573.	
PR	28-FEB-2000; 2000GB-0004695.	
XX		

PA	(CHIR ) CHIRON CORP.	
XX	(GENO-) INST GENOMIC RES.	
PI	Pizza M, Hickey E, Peterson J, Tettelin H, Venter JC, Maignani V;	
PI	Galotelli C, Mora M, Ratti G, Scarselli M, Scarliato V, Rappuoli R;	
XX	Frazer CM, Grandi G;	
DR	WPI; 2000-647603/62.	
XX		
PT	Neisseria meningitidis B full length genome sequence and open reading	
PT	frames are used to detect, treat and prevent Neisserial infections -	
XX		
PS	Claim 7; Appendix A; 692pp; English.	
XX		
CC	The present invention describes the full length genome of	
CC	Neisseria meningitidis B (NMB). The sequences in AAF21544 and AAF21607	
CC	to AAF21613 represent fragments of the NMB genomic sequence, as the	
CC	sequence was too long to go in a record on its own it was split into 8	
CC	sequences which overlap each other at the beginning and end of each	
CC	sequence by 49980 bp (i.e. the last 49980 bp of AAF21544 is repeated at	
CC	the beginning of AAF21607, the last 49980 bp of AAF21607 are repeated at	
CC	the beginning of AAF21608, and so on). AAF21545 to AAF21588 encode the	
CC	Neisseria proteins given in AAB58550 to AAB58593, and AAF21589 to	
CC	AAF21606 represent PCR primers which are used in the exemplification of	
CC	the present invention. The NMB genome and fragments from it have	
CC	antibacterial activity, and can be used in vaccines and gene therapy.	
CC	Neisseria nucleic acids, proteins and/or antibodies which binds to the	
CC	proteins can be used in compositions for treating or preventing infection	
CC	due to Neisserial bacteria or as a diagnostic reagent for detecting the	
CC	presence of Neisserial bacteria or of antibodies raised to Neisserial	
CC	bacteria. Computers, computer memory, computer storage medium or computer	
CC	databases can be used in a search to identify open reading frames (ORFs)	
CC	or coding sequences within the NMB genome. The DNA sequences provide	
CC	or further opportunities to find antigenic or immunogenic proteins which are	
CC	more effective in vaccines than the outer membrane proteins currently	
CC	used.	
XX		
SQ	Sequence 349980 BP; 86473 A; 95646 C; 85908 G; 81953 T; 0 other;	
	Query Match 26.0%; Score 834; DB 21; Length 349980;	
	Best Local Similarity 99.4%; Pred. No. 0;	
	Matches 1234; Conservative 0; Mismatches 8; Indels 0; Gaps 0;	
QY	1963 CTGCTGGCTTCCTCGACAGCGTCGAAAAACACAGCGGGCAGTGAAAGCGCACAGCTGTCC	2022
Db	263419 CTGCTGGCTTCCTCGACAGCGTCGAAAAACACAGCGGGCAGTGAAAGCGCACAGCTGTCC	263478
QY	2023 TATTATGTCGTCGCGGCAATCGCGCAGCGACTGCTTCGGCAGCGGCACATTCGCGGCC	2082
Db	263479 TATTATGTCGTCGCGGCAATCGCGCAGCGACTGCTTCGGCAGCGGCACATTCGCGGCC	263538
QY	2083 GCGGCTGTAACACAGCCCTAGAACAGGGCGCAGCAATCTGGAACACCTGATGTCGAA	2142
Db	263539 GCGGCTGTAACACAGCCCTAGAACAGGGCGCAGCAATCTGGAACACCTGATGTCGAA	263598
QY	2143 CTGGATGCTCCGAAATCATCCGCAACACCCGAGAGCGTTGAAACTCGCGGCCCGCACCC	2202
Db	263599 CTGGATGCTCCGAAATCATCCGCAACACCCGAGAGCGTTGAAACTCGCGGCCCGCACCC	263658
QY	2203 ACAGATATGCGGGGATCGCCCTTACGGCGCAACTTTCCGGCGCAGCGGCGGTACAG	2262
Db	263659 ACAGATATGCGGGGATCGCCCTTACGGCGCAACTTTCCGGCGCAGCGGCGGTACAG	263718
QY	2263 CATGCGAATGCGCGCAGCGGTGTACGCATCTTCAACAGTCTCGCGCTACCTCTATGCC	2322
Db	263719 CATGCGAATGCGCGCAGCGGTGTACGCATCTTCAACAGTCTCGCGCTACCTCTATGCC	263778
QY	2323 GACAGTACCGCGCCCATGCGGATATGCGGAGCGCGCGCTGAAAGCCGTATCGACGGG	2382
Db	263779 GACAGTACCGCGCCCATGCGGATATGCGGAGCGCGCGCTGAAAGCCGTATCGACGGG	263838
QY	2383 TTGGACCAACAGCTACGGGTCTGGCGTCTATCGGCAACCAACAGGACGGTGAACG	2442



Db 263839 TTGGACCAACACGGGCACGGGTCTGCGGTCATCGCGCAAAACCCAAACAGACAGCGTGGAAACG 263898  
Qy 2443 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGCAGTACCACAAACCGTGGCATTTGCC 2502  
|||||  
Db 263899 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGCAGTACCACAAACCGTGGCATTTGCC 263958  
Qy 2503 GGGAAACAGGGCGGAAATATACGACAGACGCGCCACACTGGGGCATGGACACACACATGG 2562  
|||||  
Db 263959 GCGAAACAGGGCGGAAATATACGACAGACGCGCCACACTGGGCATGGACGACGACATGG 264018  
Qy 2563 AGCGAAACAGTGAATGCAAAACCCACACACATTAAGTCTGTTTGGAGGCATACGGCAC 2622  
|||||  
Db 264019 AGCGAAACAGTGAATGCAAAACCCACACACATTAAGTCTGTTTGGAGGCATACGGCAC 264078  
Qy 2623 GATGCGGGCGATATCGGCTATCTCAAAGCCCTGTCTCTACGGACGCTACAAAACAGC 2682  
|||||  
Db 264079 GATGCGGGCGATATCGGCTATCTCAAAGCCCTGTCTCTACGGACGCTACAAAACAGC 264138  
Qy 2683 ATCAGCCGACAGCAGCGGTGCGGACGAACATGCGGAAGGACGCGTCAACGGCACGCTGATG 2742  
|||||  
Db 264139 ATCAGCCGACAGCAGCGGTGCGGACGAACATGCGGAAGGACGCGTCAACGGCACGCTGATG 264198  
Qy 2743 CAGCTTGGCGGCACTGGCGGTGTCAACCTTCGTTTGGCGCAACGGGAGATTGACGGTC 2802  
|||||  
Db 264199 CAGCTTGGCGGCACTGGCGGTGTCAACCTTCGTTTGGCGCAACGGGAGATTGACGGTC 264258  
Qy 2803 GAAGCGGTCTGCGCTAGCACCTGCTCAAACAGGATGCAATTCGCCGAAAAAGGCGAGTGCT 2862  
|||||  
Db 264259 GAAGCGGTCTGCGCTAGCACCTGCTCAAACAGGATGCAATTCGCCGAAAAAGGCGAGTGCT 264318  
Qy 2863 TTGGGCTGAGCGGCAACAGCCTCACTGAAGGCACACTGGTGGGACTCGCGGCTCTGAAG 2922  
|||||  
Db 264319 TTGGGCTGAGCGGCAACAGCCTCACTGAAGGCACACTGGTGGGACTCGCGGCTCTGAAG 264378  
Qy 2923 CTGTGCGCAACCTTACGCGATAAACCGTCTCTGTTTGAACGGCGGGCGTGGAAACGGCAC 2982  
|||||  
Db 264379 CTGTGCGCAACCTTACGCGATAAACCGTCTCTGTTTGAACGGCGGGCGTGGAAACGGCAC 264438  
Qy 2983 CTGAACGACGCGACTACACGCTAACGGCGGCGCTTTACCGGCGGCACTGCACAAACCGCG 3042  
|||||  
Db 264439 CTGAACGACGCGACTACACGCTAACGGCGGCGCTTTACCGGCGGCACTGCACAAACCGCG 264498  
Qy 3043 AAGACGGGGCACGCAATATGCGGCACACCCCGCTGTTGCGGCTCGGGCGCGGATGTC 3102  
|||||  
Db 264499 AAGACGGGGCACGCAATATGCGGCACACCCCGCTCTGTTGCGGCGCTGCGGCGGATGTC 264558  
Qy 3103 GAATTCGGCAACGCGTGGAAACGGCTTGGCAGCTTACAGCTACGCGCGGTTCCAAACAGTAC 3162  
|||||  
Db 264559 GAATTCGGCAACGCGTGGAAACGGCTTGGCAGCTTACAGCTACGCGCGGTTCCAAACAGTAC 264618  
Qy 3163 GGCAACCCACAGCGGACGAGTCGGCGTAGGCTACCGGTTCTGA 3204  
|||||  
Db 264619 GGCAACCCACAGCGGACGAGTCGGCGTAGGCTACCGGTTCTGA 264660

RESULT 8  
AAS43875  
ID AAS43875 standard; DNA; 3939 Bp.  
XX  
AC AAS43875;  
XX  
DT 18-DEC-2001 (first entry)  
XX  
DE Neisseria meningitidis fusion protein delta-G983-741 DNA.  
XX  
KW Neisseria gonorrhoeae; leader peptide; fusion protein; ORF46.1; ds;  
KW Neisserial protein.  
XX  
OS Neisseria meningitidis.  
XX  
PN WO200164922-A2.  
XX

PD 07-SEP-2001.  
XX  
PF 28-FEB-2001; 2001WO-IB00452.  
XX  
PR 28-FEB-2000; 2000GB-0004695.  
PR 13-NOV-2000; 2000GB-0027675.  
XX  
PA (CHIR-) CHIRON SPA.  
PI  
PI Arico MB, Comanducci M, Galeotti C, Masignani V, Guillian MM;  
PI Pizza M;  
XX  
DR WPI; 2001-582163/65.  
DR P-PSDB; AAU27576.  
XX  
PT Producing heterologous proteins from Neisseria meningitidis and N.  
PT gonorrhoeae -  
XX  
PS Example 15; Page 44-45; 119pp; English.  
XX  
CC The invention relates to methods for the heterologous expression of  
CC Neisserial proteins from Neisseria meningitidis and Neisseria  
CC gonorrhoeae. At least one domain in the protein is deleted, e.g. the  
CC leader peptide, and may be replaced by a domain from a different protein  
CC to make a fusion protein, in order to enhance heterologous expression of  
CC Neisserial proteins. Also, a region of a protein, such as a poly-glycine  
CC stretch, can be mutated to enhance expression. The proteins used in the  
CC processes include ORF46.1, 287, 741, 919, 953, 961 and 983. Sequences  
CC AAS43868-AAS43905 represent DNA molecules encoding Neisserial proteins  
CC and peptide regions of proteins of the invention.  
XX  
SQ Sequence 3939 Bp; 1046 A; 1100 C; 1107 G; 686 T; 0 other;  
  
Query Match 25.9%; Score 831; DB 22; Length 3939;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 1231; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
  
Qy 1963 CTGCTGGCTTCCCTCGACAGCGTTCGAAAAACAGCGGCACTGAAGGCGACACGCTGTCC 2022  
Db 1906 CTGCTGGCTTCCCTCGACAGCGTTCGAAAAACAGCGGCACTGAAGGCGACACGCTGTCC 1965  
Qy 2023 TATTATGTCGTCGGGCAATGCGGCACGAGTGTTCGGCAGCGGCACATTCGCGGCC 2082  
Db 1966 TATTATGTCGTCGGGCAATGCGGCACGAGTGTTCGGCAGCGGCACATTCGCGGCC 2025  
Qy 2083 GCCGTCTGAAACACGCGGTAGAACAGGCGCGCAGCAATCTGGAACCTGATGGTCGA 2142  
Db 2026 GCCGTCTGAAACACGCGGTAGAACAGGCGCGCAGCAATCTGGAACCTGATGGTCGA 2085  
Qy 2143 CTGGATGCTCCGAATCATCCGCAACACCCGAGAGCGGTGGAACCTGCGCGCCGACCGC 2202  
Db 2086 CTGGATGCTCCGAATCATCCGCAACACCCGAGAGCGGTGGAACCTGCGCGCCGACCGC 2145  
Qy 2203 ACAGATATGCGGGCATCCGCCCTACGCGCAACTTTCGGGCGACGCGGCGATACAG 2262  
Db 2146 ACAGATATGCGGGCATCCGCCCTACGCGCAACTTTCGGGCGACGCGGCGATACAG 2205  
Qy 2263 CATCGAATGCGCGGACGCGGTGTACGCACTTCAACAGCTCGCGGCTACCGTCTATGCC 2322  
Db 2206 CATCGAATGCGCGGACGCGGTGTACGCACTTCAACAGCTCGCGGCTACCGTCTATGCC 2265  
Qy 2323 GACAGTACCGCGCCCATGCGGATATGCAAGGACGCCCGGCTGAAAGCGGTATCGGACGG 2382  
Db 2266 GACAGTACCGCGCCCATGCGGATATGCAAGGACGCCCGGCTGAAAGCGGTATCGGACGG 2325  
Qy 2383 TTGACACACAGCTACGGGTCTGCGCGTCTATCGCGCAAAACCCACAGACGCGTGAACG 2442  
Db 2326 TTGACACACAGCTACGGGTCTGCGCGTCTATCGCGCAAAACCCACAGACGCGTGAACG 2385  
Qy 2443 TGGGAACAGGCGGTGTTGAAGGCAAAATGCGGCGCAGTACCCAAACCGTGGCAATGCC 2502  
Db 2386 TGGGAACAGGCGGTGTTGAAGGCAAAATGCGGCGCAGTACCCAAACCGTGGCAATGCC 2445



Qy	2503	CGGAAACCGCGGCAAAATACGACAGCGCCGCCACACTGGGCGATGGGACACAGCACATATGG	2562
Db	2446	CGGAAACCGCGGCAAAATACGACAGCGCGCCACACTGGCATGGGACGACGACACATGG	2505
Qy	2563	AGCGAAACAGTGCAAATCGAAACCGACAGCATTAGCTCTGTTTCAGGCGATACGGCAC	2622
Db	2506	AGCGAAACAGTGCAAATCGAAACCGACAGCATTAGCTCTGTTTCAGGCGATACGGCAC	2565
Qy	2523	GATCGGGGGGATATCGGCTATCTCAAGGCGCTGTTCTCTTACGACGCTACAAAAACAGC	2682
Db	2566	GATCGGGGGGATATCGGCTATCTCAAGGCGCTGTTCTCTTACGACGCTACAAAAACAGC	2625
Qy	2683	ATCAGCGCGACGACCGTGGCGAGCAATCGGAAGCAGCGCTCAAGGCGACGCTCATG	2742
Db	2626	ATCAGCGCGACGACCGTGGCGAGCAATCGGAAGCAGCGCTCAAGGCGACGCTCATG	2685
Qy	2743	CAGCTGGGGCGACTGGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTTACGGTTC	2802
Db	2686	CAGCTGGGGCGACTGGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTTACGGTTC	2745
Qy	2803	GAAGGGGTCTCGGCTACGACCTGTCTCAACAGGATGCATTCGCGGAAAAGGCGAGTGCT	2862
Db	2746	GAAGGGGTCTCGGCTACGACCTGTCTCAACAGGATGCATTCGCGGAAAAGGCGAGTGCT	2805
Qy	2863	TTGGGCTTGGAGCGGCAACAGCCTCACTCAAGGCAACACTGGTTCGAGCTCGCGGGTCTCAAG	2922
Db	2806	TTGGGCTTGGAGCGGCAACAGCCTCACTCAAGGCAACAGCCTGGTTCGAGCTCGCGGGTCTCAAG	2865
Qy	2923	CTGTGCAACCCCTTGAGCGCATTAAGCCGTCTCTGTTTGCAACGGGGGGGTGGAAACGGCAC	2982
Db	2866	CTGTGCAACCCCTTGAGCGCATTAAGCCGTCTCTGTTTGCAACGGGGGGGTGGAAACGGCAC	2925
Qy	2983	CTGAACGGACCGGACTACACGGTTAAACGGGGCGCTTTACCGGCGCGACTGCAGACAAACGGC	3042
Db	2926	CTGAACGGACCGGACTACACGGTTAAACGGGGCGCTTTACCGGCGCGACTGCAGACAAACGGC	2985
Qy	3043	AAGACGGGGCGACGCAATATGCCGACACCCCGCTTGTTGCGCGTCTGGGCGGGGATGTCT	3102
Db	2986	AAGACGGGGCGACGCAATATGCCGACACCCCGCTTGTTGCGCGTCTGGGCGGGGATGTCT	3045
Qy	3103	GAATTTCGGCAACGGCTTGAACGGCTTGGCACGTTACAGCTACGCCGTTTCCAAACAGCTAC	3162
Db	3046	GAATTTCGGCAACGGCTTGAACGGCTTGGCACGTTTACAGCTACGCCGTTTCCAAACAGCTAC	3105
Qy	3163	GGCAACACACAGCGGACGAGTCGGCGCTAGGCTACCGGTTTC	3201
Db	3106	GGCAACACACAGCGGACGAGTCGGCGCTAGGCTACCGGTTTC	3144

RESULT 9	
AA543880	
ID	AA543880 standard; DNA; 3939 BP.
XX	XX
XX	AA543880;
XX	AC
XX	XX
DT	18-DEC-2001 (first entry)
XX	XX
DE	Neisseria meningitidis fusion protein delta-G741-983 DNA.
XX	XX
KW	Neisseria gonorrhoeae; leader peptide; fusion protein; ORF46.1; ds;
KW	Neisserial protein.
XX	XX
OS	Neisseria meningitidis.
OS	Synthetic.
PN	XX
PN	WO200164922-A2.
XX	XX
PD	07-SEP-2001.
XX	XX
PF	28-FEB-2001; 2001WO-IB00452.
XX	XX
PR	28-FEB-2000; 2000GB-0004695.
PR	13-NOV-2000; 2000GB-0027675.
PR	XX

XX (CHIR-) CHIRON SPA.  
XX PA  
XX  
XX PI Arico MB, Comanducci M, Galeotti C, Masignani V, Gulliani MM;  
XX PI Pizza M;  
XX  
XX WPI: 2001-582163/65.  
XX DR P-PSDB; RAU27581.  
XX  
XX Producing heterologous proteins from *Neisseria meningitidis* and *N.*  
XX gonorrhoeae -  
XX  
XX Example 15; Page 51-52; 119pp; English.  
XX  
XX The invention relates to methods for the heterologous expression of  
XX Neisserial proteins from *Neisseria meningitidis* and *Neisseria*  
XX gonorrhoeae. At least one domain in the protein is deleted, e.g. the  
XX leader peptide, and may be replaced by a domain from a different protein  
XX to make a fusion protein, in order to enhance heterologous expression of  
XX Neisserial proteins. Also, a region of a protein, such as a poly-glycine  
XX stretch, can be mutated to enhance expression. The proteins used in the  
XX processes include ORF46.1, 287, 741, 919, 953, 961 and 983. Sequences  
XX AAS43868-AAS43905 represent DNA molecules encoding Neisserial proteins  
XX and peptide regions of proteins of the invention.  
XX  
XX Sequence 3939 BP; 1046 A; 1103 C; 1106 G; 684 T; 0 other;  
XX SQ

Query Match 25.9%; Score 831; DB 22; Length 3939;									
Best Local Similarity 99.4%; Pred. No. 0;									
Matches 1231; Conservative 0; Mismatches 8; Indels 0; Gaps 0;									
Qy	1963	CTGTGGCTTCCCTCGACAGCGTGCAGAAACACGGCGCAGTGAGGGCGACACGCTGTCC	2022						
Db	2674								
Qy	2023	TATTATGTCCGTTCGGCGCAATTCGGCAGCGACTGCTTCGGCAGCGGCACATTCCTCGCGCC	2082						
Db	2734								
Qy	2083	GCCGGTCTGAACACAGCCGCTAGACAGAGGGCGCAGCAATCTTGAAACACCTGATGGTCGAA	2142						
Db	2794								
Qy	2143	CTGGATGCCTCCGAATCATCCGCAACACCCGAGAGCGTTGAAACTTGGCGCCGCGACCGC	2202						
Db	2854								
Qy	2203	ACAGATATGCCGGGCATCCGCCCTTACGGCGCAACTTTCCGCGCAGCGGCAGCGGTACAG	2262						
Db	2914								
Qy	2263	CATGGGAATGCCGGGAGCGGTGACGCATCTTCAACAGTCTTCGCGGTACCGTCTATGCC	2322						
Db	2974								
Qy	2323	GACAGTACCGCGGCCCATTCGCGATGATCAGGAGACCGCGCTGAAAGCCGTATCGAGACGG	2382						
Db	3034								
Qy	2383	TTGGAACCAACAGCTACGGGTCTGCGGCTCATCGCGCAAAACCCAAACAGGACGGTGGAAACG	2442						
Db	3094								
Qy	2443	TGGAACCAACAGCGCACGGGTCTGCGGCTCATTCGGCGCAAAACCCAAACAGGACGGTGGAAACG	3153						
Db	3154								
Qy	2503	CGGAAACCGGCGAAAATACGACAGACGCCGCCACACTTGGGCATGGGACACAGCACATGG	2562						
Db	3214								
Qy	2563	AGCGAAAACAGTGCAAAATGCAAAAACCGACAGCAATTAGTCTGTTGCAAGGCATACGGCAC	2622						
Db	3273								

Db 3274 ACGAAAAACAGTGCATAATCAAAAACCCAGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 3333  
Qy 2623 GATGCGGGCGATATCGGCTATCTCAAAGCCCTGTTCTCTACGAGCGCTACAAAACACG 2682  
Db 3334 GATGCGGGCGATATCGGCTATCTCAAAGCCCTGTTCTCTACGAGCGCTACAAAACACG 3393  
Qy 2683 ATCAGCCGCGACACCGGTGCGGACGAACATGCGGAAGGAGGCTCAACGGCACGCTGATG 2742  
Db 3394 ATCAGCCGCGACACCGGTGCGGACGAACATGCGGAAGGAGGCTCAACGGCACGCTGATG 3453  
Qy 2743 CAGCTGGGCGCACTGGCGGCTGTCAACGTTCCGTTCCGCGCAACGGGAGATTGACGGTC 2802  
Db 3454 CAGCTGGGCGCACTGGCGGCTGTCAACGTTCCGTTCCGCGCAACGGGAGATTGACGGTC 3513  
Qy 2803 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCAATTCGCCGAAAAGCGAGTGT 2862  
Db 3514 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCAATTCGCCGAAAAGCGAGTGT 3573  
Qy 2863 TTGGGCTGAGGGGCAACAGCCTCACTGAAGGCACACTGGTGGACTCGGGGTCTTGAAG 2922  
Db 3574 TTGGGCTGAGGGGCAACAGCCTCACTGAAGGCACACTGGTGGACTCGGGGTCTTGAAG 3633  
Qy 2923 CTGTCCACACCTTGAGCGATTAAGCCGTCTCTTTTGCAACGGCGGGCTGGAACCGGAC 2982  
Db 3634 CTGTCCACACCTTGAGCGATTAAGCCGTCTCTTTTGCAACGGCGGGCTGGAACCGGAC 3693  
Qy 2983 CTGAACGAGCGACTACACGCTTAACGGCGGCTTTACCGCGCGACTGCAGCAACCGG 3042  
Db 3694 CTGAACGAGCGACTACACGCTTAACGGCGGCTTTACCGCGCGACTGCAGCAACCGG 3753  
Qy 3043 AAGACGGGGCGACGAATATGCCGCGACACCCCGCTGTTGCGGGTCTGGGCGCGGATGTC 3102  
Db 3754 AAGACGGGGCGACGAATATGCCGCGACACCCCGCTGTTGCGGGTCTGGGCGCGGATGTC 3813  
Qy 3103 GAATTCGGCAACGGCTGGCAACGGCTTGGCACGCTTACAGCTAGCCGGTTCCAAACAGTAC 3162  
Db 3814 GAATTCGGCAACGGCTGGCAACGGCTTGGCACGCTTACAGCTAGCCGGTTCCAAACAGTAC 3873  
Qy 3163 GGCACACACAGCGGACGAGTCTGGCGGTAGCTACCCGGTTC 3201  
Db 3874 GGCACACACAGCGGACGAGTCTGGCGGTAGCTACCCGGTTC 3912

RESULT 10  
AAD17040  
ID AAD17040 standard; DNA; 3939 BP.  
XX  
AC AAD17040;  
XX  
DT 29-NOV-2001 (first entry)  
XX  
DE N. meningitidis strain 2996 delta G983-741 fusion DNA.  
XX  
KW Heterologous expression; Neisserial protein;  
KW delta G983-741 fusion protein; ds.  
XX  
OS Neisseria meningitidis 2996.  
XX  
FH Key Location/Qualifiers  
CDS 1..3939  
FT /\*tag= a  
FT /product= "N. meningitidis strain 2996 delta  
FT G983-741 fusion protein"  
XX  
PN WO200164920-A2.  
XX  
PD 07-SEP-2001.  
XX  
PF 28-FEB-2001; 2001WO-IB00420.  
XX  
PR 28-FEB-2000; 2000GB-0004695.  
PR 13-NOV-2000; 2000GB-0027675.

PA (CHIR-) CHIRON SPA.  
XX  
PI Arico MB, Comanducci M, Galeotti C, Masignani V, Giuliani MM;  
PI Pizza M;  
XX  
DR WPI; 2001-557776/62.  
DR P-PSDB; AAE10023.  
XX  
PT Heterologous expression for the expression of two or more Neisserial  
PT proteins in fused state  
XX  
XX Example 3; Page 15-16; 52pp; English.  
XX  
CC The present invention relates to a method for simultaneous heterologous  
CC expression of two or more Neisserial proteins which are in a fused  
CC state. The method is useful for simultaneous heterologous expression of  
CC two or more Neisserial proteins. A protein that may be unstable or  
CC poorly expressed on its own is assisted by adding a suitable hybrid  
CC partner and commercial manufacture is simplified-only one expression and  
CC purification need to be employed in order to produce two separately-  
CC useful proteins. The present sequence is a DNA encoding  
CC Neisseria meningitidis (serogroup B, strain 2996) delta G983-741  
CC fusion protein.  
XX  
SQ Sequence 3939 BP; 1046 A; 1100 C; 1107 G; 686 T; 0 other;  
Query Match 25.9%; Score 831; DB 22; Length 3939;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 1231; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
Qy 1963 CTGCTGGCTTCCTCGACAGCGTCGAAAAACACGCGGCGAGTGAAGGCACACGCTGTCC 2022  
Db 1906 CTGCTGGCTTCCTCGACAGCGTCGAAAAACACGCGGCGAGTGAAGGCACACGCTGTCC 1965  
Qy 2023 TATTATGTCCTCGCGGCAATCGGCGACGAGTCTTCCGCGAGCGCAATTCGCGGCC 2082  
Db 1966 TATTATGTCCTCGCGGCAATCGGCGACGAGTCTTCCGCGAGCGCAATTCGCGGCC 2025  
Qy 2083 GCCGTGCTGAAACACGCGCTAGAACAGGCGCGCAGCAATCTGAAAAACCTGATGTCGAA 2142  
Db 2026 GCCGTGCTGAAACACGCGCTAGAACAGGCGCGCAGCAATCTGAAAAACCTGATGTCGAA 2085  
Qy 2143 CTGGATGCTCCGAAATCATCCGCAACACCCGAGACGGTTGAAACTCGCGGCCCGCACCCG 2202  
Db 2086 CTGGATGCTCCGAAATCATCCGCAACACCCGAGACGGTTGAAACTCGCGCACCGCACCGC 2145  
Qy 2203 ACAGATATGCGGGGATCCGCCCTACGCGGCAACTTTCGCGCGACGCGCGCTAGACG 2262  
Db 2146 ACAGATATGCGGGGATCCGCCCTACGCGGCAACTTTCGCGCGACGCGCGCTAGACG 2205  
Qy 2263 CATGCGAATGCCGCGACGCTGTACGCACTTCAACAGTCTCGCCGCTACCGTCTATGCC 2322  
Db 2206 CATGCGAATGCCGCGACGCTGTACGCACTTCAACAGTCTCGCCGCTACCGTCTATGCC 2265  
Qy 2323 GACAGTACCGCGCCCATCCGATATGACGGACCGCGGCTGAAAGCCGCTATCGGACGG 2382  
Db 2266 GACAGTACCGCGCCCATCCGATATGACGGACCGCGGCTGAAAGCCGCTATCGGACGG 2325  
Qy 2383 TTGGACCAACAGCTACGGGTCTGCGGTATCGCGCAAAACCAACAGACAGCGTGAAGC 2442  
Db 2326 TTGGACCAACAGCGGCTCTGCGGTATCGCGCAAAACCAACAGACAGCGTGAAGC 2385  
Qy 2443 TGGGAACAGGCGGTCTTGAAGGCAAAATGCGGGCAGTACCCAAACCTCGGCATTGCC 2502  
Db 2386 TGGGAACAGGCGGTCTTGAAGGCAAAATGCGGGCAGTACCCAAACCTCGGCATTGCC 2445  
Qy 2503 GCGAAAAACCGGAAAAATACGACAGCGCCGACACTGGGCATGGGACACACACATGG 2562  
Db 2446 GCGAAAAACCGGAAAAATACGACAGCGCCGACACTGGGCATGGGACGACACATGG 2505  
Qy 2563 AGCGAAAAACAGTGCRAAATGCAAAAACCGACAGCAATTAGTCTGTTTCAGGCATACGGCAC 2622  
Db 2506 AGCGAAAAACAGTGCRAAATGCAAAAACCGACAGCAATTAGTCTGTTTCAGGCATACGGCAC 2565





Db 3625 ATCAGCGCAGCACCCTGCGGAGCAACATGCGAAGAGCGCTCAACGACGCTGATG 3684  
QY 2743 CAGCTGGCGCACTGGCGCGTGTCAACGTTCCGTTTCCGCCAACAGCGGAGATTTGACGGTC 2802  
Db 3685 CAGCTGGCGCACTGGCGCGTGTCAACGTTCCGTTTCCGCCAACAGCGGAGATTTGACGGTC 3744  
QY 2803 GAAGGGCGTCTGGCTACGACCTGCTCAACAGAGATGCTATTCGCCGAAAAAGCAGTGCT 2862  
Db 3745 GAAGGGCGTCTGGCTACGACCTGCTCAACAGAGATGCTATTCGCCGAAAAAGCAGTGCT 3804  
QY 2863 TTGGCTGGAGCGGCACACGCTCACTGAAGGCACACTGTCGCGACTCGCGGCTCTGAAG 2922  
Db 3805 TTGGCTGGAGCGGCACACGCTCACTGAAGGCACACTGTCGCGACTCGCGGCTCTGAAG 3864  
QY 2923 CTGTGCGCAACCCCTGAGCGATAAAGCCGCTCTGTTTGAACGGCGGCGTGGAAACGCGAC 2982  
Db 3865 CTGTGCGCAACCCCTGAGCGATAAAGCCGCTCTGTTTGAACGGCGGCGTGGAAACGCGAC 3924  
QY 2983 CTGAACGGACGCGACTACAGCGTAACGGGCGGCTTTTACCGGCGGACTGACGCAACCCGGC 3042  
Db 3925 CTGAACGGACGCGACTACAGCGTAACGGGCGGCTTTTACCGGCGGACTGACGCAACCCGGC 3984  
QY 3043 AAGACGGGCGCAGCAATATGCCGCACACCGCGCTGTTGCCGCTGTCGGCGCGGATGTC 3102  
Db 3985 AAGACGGGCGCAGCAATATGCCGCACACCGCGCTGTTGCCGCTGTCGGCGCGGATGTC 4044  
QY 3103 GAATTTCGGCAACGGCTGGAACGGCTTGGCAGCTTACAGCTACGCGGCTTCCAAACAGTAC 3162  
Db 4045 GAATTTCGGCAACGGCTGGAACGGCTTGGCAGCTTACAGCTACGCGGCTTCCAAACAGTAC 4104  
QY 3163 GGCAACACACAGCGAGCTGGCGGTAGGCTACCGGTTTC 3201  
Db 4105 GGCAACACACAGCGAGCTGGCGGTAGGCTACCGGTTTC 4143

RESULT 13

AAD17055  
ID AAD17055 standard; DNA; 4170 BP.  
XX  
AC AAD17055;  
XX  
DT 29-NOV-2001 (first entry)  
DE N. meningitidis strain 2996 961c-983 fusion DNA.  
XX  
DE Heterologous expression; Neisserial protein;  
KW delta 961c-983 fusion protein; ds.  
KW  
XX Neisseria meningitidis 2996.  
OS  
XX  
FH Key Location/Qualifiers  
FT CDS 1..4170  
FT /tag= a  
FT /product= "N. meningitidis strain 2996 961c-983  
fusion protein"  
FT  
XX  
XX WO200164920-A2.  
XX  
XX  
PD 07-SEP-2001.  
XX  
XX 28-FEB-2001; 2001WO-IB00420.  
XX  
XX 28-FEB-2000; 2000GB-0004695.  
PR 13-NOV-2000; 2000GB-0027675.  
XX  
XX (CHIR-) CHIRON SPA.  
PA  
XX Arico MB, Comanducci M, Galeotti C, Masignani V, Giuliani MM;  
PI Pizza M;  
PI  
XX WPI; 2001-557776/62.  
DR P-PSDB; AAE10040.  
XX  
XX

PT Heterologous expression for the expression of two or more Neisserial  
PT proteins in fused state  
XX  
PS Example 23; Page 31; 52pp; English.  
XX  
CC The present invention relates to a method for simultaneous heterologous  
CC expression of two or more Neisserial proteins which are in a fused  
CC state. The method is useful for simultaneous heterologous expression of  
CC two or more Neisserial proteins. A protein that may be unstable or  
CC poorly expressed on its own is assisted by adding a suitable hybrid  
CC partner and commercial manufacture is simplified-only one expression and  
CC purification need to be employed in order to produce two separately-  
CC useful proteins. The present sequence is a DNA encoding  
CC Neisseria meningitidis (serogroup B, strain 2996) 961c-983 fusion  
CC protein.  
XX  
SQ Sequence 4170 BP; 1185 A; 1138 C; 1119 G; 728 T; 0 other;

Query Match 25.9%; Score 831; DB 22; Length 4170;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 1231; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1963 CTGCTGCTTCCCTCGACAGGCTCGAAAAACACAGCGGCGAGTGAAGCGGACACGCTGTC 2022  
Db 2905 CTGCTGCTTCCCTCGACAGGCTCGAAAAACACAGCGGCGAGTGAAGCGGACACGCTGTC 2964  
QY 2023 TATTATCTCGTTCGCGCAATGCGGCAGGACTGCTTTCGGCAGCGGCACATTTCGCGCGCC 2082  
Db 2965 TATTATCTCGTTCGCGCAATGCGGCAGGACTGCTTTCGGCAGCGGCACATTTCGCGCGCC 3024  
QY 2083 GCCGCTCTGAAACACGCGCTAGAACAGGCGGCGAGCAATCTGAAAAACCTGATGGTCGA 2142  
Db 3025 GCCGCTCTGAAACACGCGCTAGAACAGGCGGCGAGCAATCTGAAAAACCTGATGGTCGA 3084  
QY 2143 CTGATGCTTCCGCAATCATCGGCAACACCGAGAGCGTTGAAACTGCGGCGCGCGACCGC 2202  
Db 3085 CTGATGCTTCCGCAATCATCGGCAACACCGAGAGCGTTGAAACTGCGGCGCGCGACCGC 3144  
QY 2203 ACAGATATCGCGGCATCCGCCCTACGCGCAACTTTCGCGCAGCGGCGAGCGGTACAG 2262  
Db 3145 ACAGATATCGCGGCATCCGCCCTACGCGCAACTTTCGCGCAGCGGCGAGCGGTACAG 3204  
QY 2263 CATGCGAATGCGCGCGAGCGGTGTACGCATCTTCAACAGTCTCGCGCGTACCGTCTATGCC 2322  
Db 3205 CATGCGAATGCGCGCGAGCGGTGTACGCATCTTCAACAGTCTCGCGCGTACCGTCTATGCC 3264  
QY 2323 GACAGTACCGCGCCCATGCCGATATGACGAGCGCGGCTGAAAGCGGTATCGGAGCGG 2382  
Db 3265 GACAGTACCGCGCCCATGCCGATATGACGAGCGCGGCTGAAAGCGGTATCGGAGCGG 3324  
QY 2383 TTGGACCAACACGCTACGGGTCTCGCGCTCATCGCGCAAAACCCACAGGAGCGTGGAAACG 2442  
Db 3325 TTGGACCAACACGCGACGGGTCTCGCGCTCATCGCGCAAAACCCACAGGAGCGTGGAAACG 3384  
QY 2443 TGGGAACAGGCGGTGTGAAAGGCAAAATGCGCGCGAGTAGTACCCAAACCGTCGGCATGTC 2502  
Db 3385 TGGGAACAGGCGGTGTGAAAGGCAAAATGCGCGCGAGTAGTACCCAAACCGTCGGCATGTC 3444  
QY 2503 GCGAAAAACCGCGAAATACGACAGCAGCCCGCCACACTGGGCATGGGACACAGCACATGG 2562  
Db 3445 GCGAAAAACCGCGAAATACGACAGCAGCCCGCCACACTGGGCATGGGACACAGCACATGG 3504  
QY 2563 AGCGAAAAACAGTCAAAATGCAAAACCGACAGCAATTAGTCTGTTTGGAGGCAATACGGCAG 2622  
Db 3505 AGCGAAAAACAGTCAAAATGCAAAACCGACAGCAATTAGTCTGTTTGGAGGCAATACGGCAG 3564  
QY 2623 GATCGGGCGCATATCGGCTATCTCAAAAGGCGCTGTTCTCTACGAGCGCTACAAAAACAGC 2682  
Db 3565 GATCGGGCGCATATCGGCTATCTCAAAAGGCGCTGTTCTCTACGAGCGCTACAAAAACAGC 3624  
QY 2683 ATCAGCGCAGCACCGGTGCGGAGCAACATCGGGAAGCAGCGTCAACGGCACGCTGATG 2742  
Db 3625 ATCAGCGCAGCACCGGTGCGGAGCAACATCGGGAAGCAGCGTCAACGGCACGCTGATG 3684

QY 2743 CAGCTGGCGCACTGGCGGTGTCAACGTTCCGTTTCCGCAACGGAGATTGACGGTC 2802  
|||||  
Db 3685 CAGCTGGCGCACTGGCGGTGTCAACGTTCCGTTTCCGCAACGGAGATTGACGGTC 3744  
|||||  
QY 2803 GAAGCGGTCTCGGTACGACCTGCTCAAAAGGATGCAATTCGCCGAAAGGCGATGCT 2862  
|||||  
Db 3745 GAAGCGGTCTCGGTACGACCTGCTCAAAAGGATGCAATTCGCCGAAAGGCGATGCT 3804  
|||||  
QY 2863 TTGGGCTGAGCGGCAACAGCCTCACTGAAGCAGCAGCTGTCGGACTCGCGGTCTGAAG 2922  
|||||  
Db 3805 TTGGGCTGAGCGGCAACAGCCTCACTGAAGCAGCAGCTGTCGGACTCGCGGTCTGAAG 3864  
|||||  
QY 2923 CTGTCGCAACCTTGAAGCATAAAGCCGTCCTGTTTGCAACGGCGGCGTGGAAACGCGAC 2982  
|||||  
Db 3865 CTGTCGCAACCTTGAAGCATAAAGCCGTCCTGTTTGCAACGGCGGCGTGGAAACGCGAC 3924  
|||||  
QY 2983 CTGAACGGACGCACTACAGGTAAAGCGCGCTTTACCGGCGCACTGCACCAACCGCG 3042  
|||||  
Db 3925 CTGAACGGACGCACTACAGGTAAAGCGCGCTTTACCGGCGCACTGCACCAACCGCG 3984  
|||||  
QY 3043 AAGACGGGGCAGCAATATGCCGACACCCGCTGTTGCCGCTGTCGGCGGATGTC 3102  
|||||  
Db 3985 AAGACGGGGCAGCAATATGCCGACACCCGCTGTTGCCGCTGTCGGCGGATGTC 4044  
|||||  
QY 3103 GAATTCGGCAACGGCTTGAACGGCTTGGCACGTTACAGCTACGCGGTTTCCAAACAGTAC 3162  
|||||  
Db 4045 GAATTCGGCAACGGCTTGAACGGCTTGGCACGTTACAGCTACGCGGTTTCCAAACAGTAC 4104  
|||||  
QY 3163 GGAACCAACAGCGGAGCTGCGGCTAGGCTACCGGTTCC 3201  
|||||  
Db 4105 GGAACCAACAGCGGAGCTGCGGCTAGGCTACCGGTTCC 4143  
|||||

RESULT 14

AA543877  
ID AAS43877 standard; DNA; 4179 BP.  
AC AAS43877;  
XX  
XX  
DT 18-DEC-2001 (first entry)  
XX  
XX  
DE Neisseria meningitidis fusion protein delta-g983-961c DNA.  
XX  
KW Neisseria gonorrhoeae; leader peptide; fusion protein; ORF46.1; ds;  
KW Neisseria protein.  
XX  
OS Neisseria meningitidis.  
OS Synthetic.  
PN WO200164922-A2.  
XX  
PD 07-SEP-2001.  
XX  
XX 28-FEB-2001; 2001WO-IB00452.  
PF  
XX  
PR 28-FEB-2000; 2000GB-0004695.  
PR 13-NOV-2000; 2000GB-0027675.  
XX  
XX  
PA (CHIR-) CHIRON SPA.  
XX  
XX Arico MB, Comanducci M, Galeotti C, Massignani V, Guiliani MM;  
PI Pizza M;  
XX  
XX WPI; 2001-582163/65.  
DR P-PSDB; AAU27578.  
XX  
XX Producing heterologous proteins from Neisseria meningitidis and N.  
PT gonorrhoeae -  
XX  
XX Example 15; Page 47-49; 119pp; English.  
PS  
XX The invention relates to methods for the heterologous expression of

CC Neisserial proteins from Neisseria meningitidis and Neisseria  
CC gonorrhoeae. At least one domain in the protein is deleted, e.g. the  
CC leader peptide, and may be replaced by a domain from a different protein  
CC to make a fusion protein, in order to enhance heterologous expression of  
CC Neisserial proteins. Also, a region of a protein, such as a poly-glycine  
CC stretch, can be mutated to enhance expression. The proteins used in the  
CC processes include ORF46.1, 287, 741, 919, 953, 961 and 983. Sequences  
CC AAS43868-AAS43905 represent DNA molecules encoding Neisserial proteins  
CC and peptide regions of the invention.

XX Sequence 4179 BP; 1187 A; 1140 C; 1121 G; 731 T; 0 other;

Query Match 25.9%; Score 831; DB 22; Length 4179;

Best Local Similarity 99.4%; Pred. No. 0;

Matches 1231; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1963 CTGCTGGCTTCCTCGACAGCGTCGAAAAAACAAGCGGGCAGTGAAGCGACACGCTGTCC 2022  
|||||  
Db 1906 CTGCTGGCTTCCTCGACAGCGTCGAAAAAACAAGCGGGCAGTGAAGCGACACGCTGTCC 1965  
|||||  
QY 2023 TATTATGTCGTCGCGCAATGCGGCAGCGACTGCTCGGCAGCGCACATTTCCGCGCC 2082  
|||||  
Db 1966 TATTATGTCGTCGCGCAATGCGGCAGCGACTGCTCGGCAGCGCACATTTCCGCGCC 2025  
|||||  
QY 2083 GCCGTCTGAAACACGCGCTAGAACAGGGCGGCAGCAATCTGGAAGAACCTGATGTCGAA 2142  
|||||  
Db 2026 GCCGTCTGAAACACGCGCTAGAACAGGGCGGCAGCAATCTGGAAGAACCTGATGTCGAA 2085  
|||||  
QY 2143 CTGGATGCTCCGAATCATCCGACACCCGAGAGCGTTGAAACTGCGGCGCCGCGCACCG 2202  
|||||  
Db 2086 CTGGATGCTCCGAATCATCCGACACCCGAGAGCGTTGAAACTGCGGCGCCGCGCACCG 2145  
|||||  
QY 2203 ACAGATATGCGGGCATCGCCCTACGCGCAACTTTCCGCGCAGCGCACCGCTACAG 2262  
|||||  
Db 2146 ACAGATATGCGGGCATCGCCCTACGCGCAACTTTCCGCGCAGCGCACCGCTACAG 2205  
|||||  
QY 2263 CATGCGAATGCGCGCAGGTTAGCGATCTTCAACAGTCTCGCGCTACCGCTATGTC 2322  
|||||  
Db 2206 CATGCGAATGCGCGCAGGTTAGCGATCTTCAACAGTCTCGCGCTACCGCTATGTC 2265  
|||||  
QY 2323 GACAGTACGCGCGCCATGCCGATATGAGGAGCGCGGCTGAAAGCCGTATCGACCGG 2382  
|||||  
Db 2266 GACAGTACGCGCGCCATGCCGATATGAGGAGCGCGGCTGAAAGCCGTATCGACCGG 2325  
|||||  
QY 2383 TTGGACCAACAACGCTACGGGCTCGCGCTATCGCGCAACCAACAGGACGTTGGAAGC 2442  
|||||  
Db 2326 TTGGACCAACAACGCTACGGGCTCGCGCTATCGCGCAACCAACAGGACGTTGGAAGC 2385  
|||||  
QY 2443 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGCAGTACCCAAACCGTCGGCATTTGCC 2502  
|||||  
Db 2386 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGCAGTACCCAAACCGTCGGCATTTGCC 2445  
|||||  
QY 2503 GCGAAACCGCGGCAAAATACGACAGCGCGCCACACTGGGGCATGGGACACAGCATGG 2562  
|||||  
Db 2446 GCGAAACCGCGGCAAAATACGACAGCGCGCCACACTGGGGCATGGGACACAGCATGG 2505  
|||||  
QY 2563 AGCGAAACAGTGCATAATGCAAAACCGACACATTAGTCTGTTTTCAGGCGCATACGGCAC 2622  
|||||  
Db 2506 AGCGAAACAGTGCATAATGCAAAACCGACACATTAGTCTGTTTTCAGGCGCATACGGCAC 2565  
|||||  
QY 2623 GATGCGGCGATATCGGCTATCTCAAAGGCGCTTCTCTCTACGGAGCGGTACAAAACAGC 2682  
|||||  
Db 2566 GATGCGGCGATATCGGCTATCTCAAAGGCGCTTCTCTCTACGGAGCGGTACAAAACAGC 2625  
|||||  
QY 2683 ATCAGCGCGACCGGTCGCGACGCAATGCGGAGGAGCGGTCAACGGCAGCGTGATG 2742  
|||||  
Db 2626 ATCAGCGCGACCGGTCGCGACGCAATGCGGAGGAGCGGTCAACGGCAGCGTGATG 2685  
|||||  
QY 2743 CAGCTGGGCGCACTGGCGGTGTCACGTTTCGTTTTCGCGCAACGGAGATTGTCGGGTC 2802  
|||||  
Db 2686 CAGCTGGGCGCACTGGCGGTGTCACGTTTCGTTTTCGCGCAACGGAGATTGTCGGGTC 2745  
|||||  
QY 2803 GAAGCGGTCTCGGCTACGACCTGCTCAAAACAGGATGATTCGCCGAAAAAGGACGTGCT 2862  
|||||

Db 2746 GAAGGGGGTCTGGCTACGACCTGCTCAAAACAGGATGCTTCGCCGAAAGGAGTGGCT 2805  
QY 2863 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTTCGACTCGCGGCTCTGAAG 2922  
Db 2806 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACCGTGGTTCGACTCGCGGCTCTGAAG 2865  
QY 2923 CTGTGCAACCCCTTGAGCGATAAAGCGTCTCTTTGCAACGCGGCGGTGGAACGCGAC 2982  
Db 2866 CTGTGCAACCCCTTGAGCGATAAAGCGTCTCTTTGCAACGCGGCGGTGGAACGCGAC 2925  
QY 2983 CTGAACGGAGCGCACTACACGGTAACGGGCGGCTTTACCGGCGGACTCGAGCAACCGGC 3042  
Db 2926 CTGAACGGAGCGCACTACACGGTAACGGGCGGCTTTACCGGCGGACTCGAGCAACCGGC 2985  
QY 3043 AAGACGGGGGCAACCAATATGCCGCGCACACCGCGCTGGTTGCCGCTCTGGGCGGGATGTC 3102  
Db 2986 AAGACGGGGGCAACCAATATGCCGCGCACACCGCGCTGGTTGCCGCTCTGGGCGGGATGTC 3045  
QY 3103 GAATTCGGCAACGGCTGGGAACGGCTTTGGCACGCTTACAGCTACCGCGTTTCCAAACAGTAC 3162  
Db 3046 GAATTCGGCAACGGCTGGGAACGGCTTTGGCACGCTTACAGCTACCGCGTTTCCAAACAGTAC 3105  
QY 3163 GGCACCCACAGCGGACGAGTCGGCGTAGGCTACCGGTTTC 3201  
Db 3106 GGCACCCACAGCGGACGAGTCGGCGTAGGCTACCGGTTTC 3144

RESULT 15

AAD17042  
ID AAD17042 standard; DNA; 4179 BP.  
XX  
AC AAD17042;  
XX  
DT 29-NOV-2001 (first entry)  
XX  
DE N. meningitidis strain 2996 delta G983-961c fusion DNA.  
XX  
KW Heterologous expression; Neisserial protein;  
KW delta G983-961c fusion protein; ds.  
OS Neisseria meningitidis 2996.  
FH Key Location/Qualifiers  
FT 1..4179  
FT /\*tag= a  
FT /product= "N. meningitidis strain 2996 delta  
FT G983-961c fusion protein"  
XX  
PN WO200164920-A2.  
XX  
PD 07-SEP-2001.  
XX  
PF 28-FEB-2001; 2001WO-IB00420.  
XX  
PR 28-FEB-2000; 2000GB-0004695.  
PR 13-NOV-2000; 2000GB-0027675.  
PA (CHIR-) CHIRON SPA.  
XX  
PI Arico MB, Comanducci M, Galeotti C, Massignani V, Giuliani MM;  
PI Pizza M;  
XX  
DR WPI: 2001-557776/62.  
DR P-PSDB; AAE10025.  
XX  
XX Heterologous expression for the expression of two or more Neisserial  
PT proteins in fused state  
XX  
PS Example 3; Page 17-18; 52pp; English.  
XX  
CC The present invention relates to a method for simultaneous heterologous  
CC expression of two or more Neisserial proteins which are in a fused

CC state. The method is useful for simultaneous heterologous expression of  
CC two or more Neisserial proteins. A protein that may be unstable or  
CC poorly expressed on its own is assisted by adding a suitable hybrid  
CC partner and commercial manufacture is simplified-only one expression and  
CC purification need to be employed in order to produce two separately-  
CC useful proteins. The present sequence is a DNA encoding  
CC Neisseria meningitidis (serogroup B, strain 2996) delta G983-961c  
CC fusion protein.  
XX  
SQ Sequence 4179 BP; 1187 A; 1140 C; 1121 G; 731 T; 0 other;

Query Match 25.9%; Score 831; DB 22; Length 4179;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 1231; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 1963 CTGCTGGCTTCCCTCGACAGCGTTCGAAAAACAGCGGGGAGTGAAGCGACACGCTGTCC 2022  
Db 1906 CTGCTGGCTTCCCTCGACAGCGTTCGAAAAACAGCGGGGAGTGAAGCGACACGCTGTCC 1965  
QY 2023 TATTATGTCTCGTCGCGCAATGCGGACGACTGCTTTCGGCAGCGGCACATTCGCGCC 2082  
Db 1966 TATTATGTCTCGTCGCGCAATGCGGACGACTGCTTTCGGCAGCGGCACATTCGCGCC 2025  
QY 2083 GCCGGTCTGAAACACACGCCGTAGAACAGCGGCGAGCAATCTGAAAAACCTGATGGTCGAA 2142  
Db 2026 GCCGGTCTGAAACACACGCCGTAGAACAGCGGCGAGCAATCTGAAAAACCTGATGGTCGAA 2085  
QY 2143 CTGGATGCCCTCCGAATCATCCGCAACACCGAGACGGTTGAAACTGGGGCGCGGACCCG 2202  
Db 2086 CTGGATGCCCTCCGAATCATCCGCAACACCGAGACGGTTGAAACTGGGGCGCGGACCCG 2145  
QY 2203 ACAGATATGCCGGGATCCGCCCTACGGGCGCAACTTTCGCGCAGCGGCGAGCGGTACAG 2262  
Db 2146 ACAGATATGCCGGGATCCGCCCTACGGGCGCAACTTTCGCGCAGCGGCGAGCGGTACAG 2205  
QY 2263 CATGCGAATGCCCGGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGCTCTATGCC 2322  
Db 2206 CATGCGAATGCCCGGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGCTCTATGCC 2265  
QY 2323 GACAGTACCGCGCCCATGCCGATATGCGAGGACGCGGCTGAAGCCGCTATCGGACGG 2382  
Db 2266 GACAGTACCGCGCCCATGCCGATATGCGAGGACGCGGCTGAAGCCGCTATCGGACGG 2325  
QY 2383 TTGGACACAAACGCTACGGGTCTGCGCTCATCGCGCAAAACCCAAACAGGACGTTGGAACG 2442  
Db 2326 TTGGACACAAACGCGACGGGTCTGCGGCTCATCGCGCAAAACCCAAACAGGACGTTGGAACG 2385  
QY 2443 TGGGAACAGGGCGGTGTGAAGGCAAAATGCGCGGAGTACCCAAACCGTTCGGCATTTGCC 2502  
Db 2386 TGGGAACAGGGCGGTGTGAAGGCAAAATGCGCGGAGTACCCAAACCGTTCGGCATTTGCC 2445  
QY 2503 GCGAAAAACCGGCAAAATAGCAGACGACGCGCCACACTGGGCGATGGGACACACGACATGG 2562  
Db 2446 GCGAAAAACCGGCAAAATAGCAGACGACGCGCCACACTGGGCGATGGGACACACGACATGG 2505  
QY 2563 AGCGAAAAACAGTGCATAATGCAAAAACCGACGACATTAGTCTGTTTCAGGACATACGCGAC 2622  
Db 2506 AGCGAAAAACAGTGCATAATGCAAAAACCGACGACATTAGTCTGTTTCAGGACATACGCGAC 2565  
QY 2623 GATGCGGGCGATATCGGCTATCTCAAAAGCGCTTCTCTACGAGCGCTACAAAAACAGC 2682  
Db 2566 GATGCGGGCGATATCGGCTATCTCAAAAGCGCTTCTCTACGAGCGCTACAAAAACAGC 2625  
QY 2683 ATCAGCGCGACGACCGGTGGGAGCAACATGCGGAGGACGCGTCAACGCGACGCTGATG 2742  
Db 2626 ATCAGCGCGACGACCGGTGGGAGCAACATGCGGAGGACGCGTCAACGCGACGCTGATG 2685  
QY 2743 CAGCTGGGCGCACTGGGCGGTGTCAACGTTTCGCGCAACCGGAGATTTGACGGTC 2802  
Db 2686 CAGCTGGGCGCACTGGGCGGTGTCAACGTTTCGCGCAACCGGAGATTTGACGGTC 2745  
QY 2803 GAAGCGCGTCTGCGCTACGACCTGCTCAACAGGATGCTTCGCCGAAAAAGGACGCTGCT 2862

Db 2746 GAAGCGGCTCTGCGCTACGACCTGCTCAACACAGGATGCATTCGCCGAAAGGACGTGCT 2805  
QY 2863 TTGGGCTGGAGCGCAACAGCCTCACTGAAGGCACACACTGCTCGGACTCGCGGCTCTGAAG 2922  
Db 2806 TTGGGCTGGAGCGCAACAGCCTCACTGAAGGCACACCTGCTCGGACTCGCGGCTCTGAAG 2865  
QY 2923 CTGTCGAACCCCTTGAGCGATGAAGCCGCTCOTGTTTTCGAACGCGGCGCTGGAACCGGAC 2982  
Db 2866 CTGTCGAACCCCTTGAGCGATGAAGCCGCTCCTGTTTTCGAACGCGGCGCTGGAACCGGAC 2925  
QY 2983 CTGAACGGAGCGGACACTACAGGTAACGGGCGGCTTTACGCGGCGGACTGCAAGCAACCGGC 3042  
Db 2926 CTGAACGGAGCGGACACTACAGGTAACGGGCGGCTTTACGCGGCGGACTGCAAGCAACCGGC 2985  
QY 3043 AAGACGGGGGCGAGCAATATGCCGACACCGCGCTGTTGCGGCTCTGGCGCGGATGTC 3102  
Db 2986 AAGACGGGGGCGAGCAATATGCCGACACCGCGCTGTTGCGGCGGCTCTGGCGCGGATGTC 3045  
QY 3103 GAATTCGGCAACGGCTTGAACGGCTTGGCACGTTACAGCTACGCCGTTTCCAAACAGTAC 3162  
Db 3046 GAATTCGGCAACGGCTTGAACGGCTTGGCACGTTACAGCTACGCCGTTTCCAAACAGTAC 3105  
QY 3163 GGCACACAGCGGAGCTGCGGCTAGCGTAGCGTACCGGTTTC 3201  
Db 3106 GGCACACAGCGGAGCTGCGGCTAGCGTAGCGTACCGGTTTC 3144

RESULT 16

AAS43899  
XX AAS43899 standard; DNA; 4335 BP.  
AC AAS43899;  
XX  
XX  
DT 18-DEC-2001 (first entry)  
XX  
DE Neisseria meningitidis fusion protein 961-983 DNA.  
XX  
XX Neisseria gonorrhoeae; leader peptide; fusion protein; ORF46.1; ds;  
KW Neisserial protein.  
XX  
XX Neisseria meningitidis.  
OS Synthetic.  
XX  
XX WO200164922-A2.  
XX  
XX 07-SEP-2001.  
XX  
XX 28-FEB-2001; 2001WO-IB00452.  
XX  
XX 28-FEB-2000; 2000GB-0004695.  
PR 13-NOV-2000; 2000GB-0027675.  
XX  
XX (CHIR-) CHIRON SPA.  
PA  
XX Arico MB, Comanducci M, Galeotti C, Massignani V, Guilianini MM;  
PI Pizza M;  
XX  
XX WPI: 2001-582163/65.  
DR P-PSDB; AAU27602.  
DR  
XX  
PT Producing heterologous proteins from Neisseria meningitidis and N.  
PT gonorrhoeae -  
XX  
XX Example 23; Page 68-69; 119pp; English.  
XX  
CC The invention relates to methods for the heterologous expression of  
CC Neisserial proteins from Neisseria meningitidis and Neisseria  
CC gonorrhoeae. At least one domain in the protein is deleted, e.g. the  
CC leader peptide, and may be replaced by a domain from a different protein  
CC to make a fusion protein, in order to enhance heterologous expression of  
CC Neisserial proteins. Also, a region of a protein, such as a poly-glycine  
CC stretch, can be mutated to enhance expression. The proteins used in the  
CC processes include ORF46.1, 287, 741, 919, 953, 961 and 983. Sequences

CC AAS43868-AAS43905 represent DNA molecules encoding Neisserial proteins  
CC and peptide regions of proteins of the invention.  
XX  
SQ Sequence 4335 BP; 1217 A; 1189 C; 1165 G; 764 T; 0 other;  
Query Match 25.9%; Score 831; DB 22; Length 4335;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 1231; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 1963 CTGCTGGCTTCCTCGACAGCGTCGAAAAACAGCGGGCAGTGAAGGCACACGCTGTGCC 2022  
Db 3070 CTGCTGGCTTCCTCGACAGCGTCGAAAAACAGCGGGCAGTGAAGGCACACGCTGTGCC 3129  
QY 2023 TATTATGTCCTGCGGGCAATGGGCACGACTGCTTCGGCAGCGCACATTCGCGGCC 2082  
Db 3130 TATTATGTCCTGCGGGCAATGGGCACGACTGCTTCGGCAGCGCACATTCGCGGCC 3189  
QY 2083 GCCGCTCTGAAACACGCCGTAGAACAGGCGCGCAACTCTGAAAACTGATGTCGAA 2142  
Db 3190 GCCGCTCTGAAACACGCCGTAGAACAGGCGCGCAACTCTGAAAACTGATGTCGAA 3249  
QY 2143 CTGGATGCTCCGATCATCCGACACACCCGAGACGGTTGAAACTGCGGCCCGCACCGC 2202  
Db 3250 CTGGATGCTCCGATCATCCGACACACCCGAGACGGTTGAAACTGCGGCCCGCACCGC 3309  
QY 2203 ACAGATATGCGGGGATCGCCCTACGGCGCAACTTTCGCGCGACGGCGCGCTACAG 2262  
Db 3310 ACAGATATGCGGGGATCGCCCTACGGCGCAACTTTCGCGCGACGGCGCGCTACAG 3369  
QY 2263 CATGGAATGCGCGGACGCTGACGATCTTCAACAGTCTCGCGGCTACCGCTATATGCC 2322  
Db 3370 CATGGAATGCGCGGACGCTGACGATCTTCAACAGTCTCGCGGCTACCGCTATATGCC 3429  
QY 2323 GACAGTACGCGCGCCATGCCGATATGACGAGACGCGGCTGAAAGCCGTATCGACGGG 2382  
Db 3430 GACAGTACGCGCGCCATGCCGATATGACGAGACGCGGCTGAAAGCCGTATCGACGGG 3489  
QY 2383 TTGGACCAACAGCTACGGGCTCGCGGTCATCGCGCAACCCACAGACGCTGGAAGC 2442  
Db 3490 TTGGACCAACAGCTACGGGCTCGCGGTCATCGCGCAACCCACAGACGCTGGAAGC 3549  
QY 2443 TGGGAACAGGCGGTGTTGAAGGCAAAATGCGGCGAGTACCCAAACCTCGCGCATTTGCC 2502  
Db 3550 TGGGAACAGGCGGTGTTGAAGGCAAAATGCGGCGAGTACCCAAACCTCGCGCATTTGCC 3609  
QY 2503 GCGAAACCGGCGAAATATAGACAGACGCGCCACACATGCGGCGATGGGACACACATGG 2562  
Db 3610 GCGAAACCGGCGAAATATAGACAGACGCGCCACACATGCGGCGATGGGACACACATGG 3669  
QY 2563 AGCGAAACAGTGCAAAATGCAAAACCGACAGCATTAGTCTGTTTTCAGGCGATACGGAC 2622  
Db 3670 AGCGAAACAGTGCAAAATGCAAAACCGACAGCATTAGTCTGTTTTCAGGCGATACGGAC 3729  
QY 2623 GATCGGGCGATATCGGCTATCTCAAAGCCCTGTTCTCTCTACGGACGCTTACAAAAACAGC 2682  
Db 3730 GATCGGGCGATATCGGCTATCTCAAAGCCCTGTTCTCTCTACGGACGCTTACAAAAACAGC 3789  
QY 2683 ATCAGCCGCGACCGGTGCGGACGCAACATGCGGAAGCGAGCGCTCAACGGCACGCTGATG 2742  
Db 3790 ATCAGCCGCGACCGGTGCGGACGCAACATGCGGAAGCGAGCGCTCAACGGCACGCTGATG 3849  
QY 2743 CAGCTGGGCGCACTGGGCGGTGTCAACGTTTCGCGCAACGGGAGATTGACGGTC 2802  
Db 3850 CAGCTGGGCGCACTGGGCGGTGTCAACGTTTCGCGCAACGGGAGATTGACGGTC 3909  
QY 2803 GAAGCGGTCTGCGCTACGACCTGCTCAAACAGGATGATTCGCCGAAAAAGGCACTGCT 2862  
Db 3910 GAAGCGGTCTGCGCTACGACCTGCTCAAACAGGATGATTCGCCGAAAAAGGCACTGCT 3969  
QY 2863 TTGGGCTGGAGCGGCAACAGCCCTCACTGAAGGCACACTGTTGTCGGACTCGCGGGTCTGAAG 2922  
Db 3970 TTGGGCTGGAGCGGCAACAGCCCTCACTGAAGGCACGCTGGTTCGGACTCGCGGGTCTGAAG 4029



Qy	2923	CTGTGCGCAACCCTTGACGGATAAAGCCGTCCTCTGTTTGCAACGCGGCGCTGGAAACGCGAC	2982
Db	4030	CTGTGCGCAACCCTTGACGGATAAAGCCGTCCTCTGTTTGCAACGCGGCGCTGGAAACGCGAC	4089
Qy	2983	CTGAACGGAGCGGACTACACGGTAAACGGCGGCTTTACGGCGCGGACTGCAGCAACCGGC	3042
Db	4090	CTGAACGGAGCGGACTACACGGTAAACGGCGGCTTTACGGCGCGGACTGCAGCAACCGGC	4149
Qy	3043	AAGACGGGGCACGCAATATGCCGCACACCCGCTGGTTCGCCGCTCTGGGCGCGGATGTC	3102
Db	4150	AAGACGGGGCACGCAATATGCCGCACACCCGCTGGTTCGCCGCTCTGGGCGCGGATGTC	4209
Qy	3103	GAATTCGGCAACGCGCTGGAGCGCTTGCGACGTTTACAGCTACGCCGGTTCCTCAACAGTAC	3162
Db	4210	GAATTCGGCAACGCGCTGGAGCGCTTGCGACGTTTACAGCTACGCCGGTTCCTCAACAGTAC	4269
Qy	3163	GGCAACACACGCGACGAGTCGGCGTAGGCTACCGGPTC	3201
Db	4270	GGCAACACACGCGACGAGTCGGCGTAGGCTACCGGPTC	4308
RESULT	17		
AAAD17052			
ID	AAAD17052	standard; DNA; 4335 BP.	
XX	AC	AC	
XX	AAAD17052;		
DT	29-NOV-2001	(first entry)	
XX			
DE	N. meningitidis strain 2996	961-983 fusion DNA.	
XX			
KW	Heterologous expression; Neisserial protein;	961-983 fusion protein; ds.	
OS	Neisseria meningitidis	2996.	
XX			
EH	Key	Location/Qualifiers	
FT	CDS	1..4335	
FT		/*tag= a	
FT		/product= "N. meningitidis strain 2996	961-983
FT		fusion protein"	
XX			
PN	W0200164920-A2.		
XX			
PD	07-SEP-2001.		
XX			
PF	28-FEB-2001; 2001WO-IB00420.		
XX			
PR	28-FEB-2000; 2000GB-0004695.		
PR	13-NOV-2000; 2000GB-0027675.		
XX			
PA	(CHIR-) CHIRON SPA.		
XX			
PI	Arico MB, Comanducci M, Galeotti C, Masignani V, Giuliani MM;		
PI	Pizza M;		
XX			
DR	WPI; 2001-557776/62.		
DR	P-PSDB; AAE10037.		
XX			
PT	Heterologous expression for the expression of two or more Neisserial		
PT	proteins in fused state		
XX			
PS	Example 23; Page 28-29; 52pp; English.		
XX			
CC	The present invention relates to a method for simultaneous heterologous		
CC	expression of two or more Neisserial proteins which are in a fused		
CC	state. The method is useful for simultaneous heterologous expression of		
CC	two or more Neisserial proteins. A protein that may be unstable or		
CC	poorly expressed on its own is assisted by adding a suitable hybrid		
CC	partner and commercial manufacture is simplified-only one expression and		
CC	purification need to be employed in order to produce two separately-		
CC	useful proteins. The present sequence is a DNA encoding		
CC	Neisseria meningitidis (serogroup B, strain 2996) 961-983 fusion		
CC	protein.		

Db	4030	CTGTCCGAACCCCTGAGCGATAAAGCCGTCCTGTTTTCGAACGGCGGGCGTGAACGCGAC	4089
QY	2983	CTGAACGGACGGACTACACGCTAACGGCGCGCTTTACCGGGCCGACTGCACGAACCGGC	3042
Db	4090	CTGAACGGACGGACTACACGCTAACGGCGCGCTTTACCGGGCCGACTGCACGAACCGGC	4149
QY	3043	AAAGCGGGGGACGCAATATGCCGACACCCGCCCTGGTTGCGCGCTCTGGCGCGGATGTC	3102
Db	4150	AAAGCGGGGGACGCAATATGCCGACACCCGCCCTGGTTGCGCGCTCTGGCGCGGATGTC	4209
QY	3103	GAATTCGGCAACGGCTGGNAACGGCTTGGCACGTTACAGCTACGCGCGTTCCAAACAGTAC	3162
Db	4210	GAATTCGGCAACGGCTGGNAACGGCTTGGCACGTTACAGCTACGCGCGTTCCAAACAGTAC	4269
QY	3163	GGCAACACAGCGGACGAGTCGGCGCTAGGCTACCGGTTCC	3201
Db	4270	GGCAACACAGCGGACGAGTCGGCGCTAGGCTACCGGTTCC	4308
RESULT 18			
AAS43876			
ID	AAS43876	standard; DNA; 4344 BP.	
XX	AAS43876;		
AC	AAS43876;		
XX			
DT	18-DEC-2001	(first entry)	
XX			
DE	Neisseria meningitidis	fusion protein delta-G983-961 DNA.	
XX			
KW	Neisseria gonorrhoeae;	leader peptide; fusion protein; ORF46.1; ds;	
XX			
KW	Neisserial protein.		
XX			
OS	Neisseria meningitidis.		
OS	Synthetic.		
XX			
PN	WO200164922-A2.		
XX			
PD	07-SEP-2001.		
XX			
PF	28-FEB-2001;	2001WO-IB00452.	
XX			
PR	28-FEB-2000;	2000GB-0004695.	
PR	13-NOV-2000;	2000GB-0027675.	
XX			
PA	(CHIR-) CHIRON SPA.		
XX			
PI	Arico MB, Comanducci M, Galeotti C, Masignani V, Guilianani MM;		
PI	Pizza M;		
XX			
DR	WPI; 2001-582163/65.		
DR	P-PSDB; AAU27577.		
XX			
PT	Producing heterologous proteins from Neisseria meningitidis and N. gonorrhoeae -		
PT			
PS	Example 15; Page 46-47; 119pp; English.		
XX			
CC	The invention relates to methods for the heterologous expression of		
CC	Neisserial proteins from Neisseria meningitidis and Neisseria		
CC	gonorrhoeae. At least one domain in the protein is deleted, e.g. the		
CC	leader peptide, and may be replaced by a domain from a different protein		
CC	to make a fusion protein, in order to enhance heterologous expression of		
CC	Neisserial proteins. Also, a region of a protein, such as a poly-glycine		
CC	stretch, can be mutated to enhance expression. The proteins used in the		
CC	processes include ORF46.1, 287, 741, 919, 953, 961 and 983. Sequences		
CC	AAS43868-AAS43905 represent DNA molecules encoding Neisserial proteins		
CC	and peptide regions of proteins of the invention.		
XX			
SQ	Sequence 4344 BP; 1219 A; 1167 G; 767 T; 0 other;		
Query Match 25.9%; Score 831; DB 22; Length 4344;			
Best Local Similarity 99.4%; Pred. No. 0;			
Matches 1231; Conservative 0; Mismatches 8; Indels 0; Gaps 0;			

QY	1963	CTGCTGGCTTCCCTCGACAGCGTTCGAAAAACAGCGGCGAGTGAAGGCGACACGCTGCC	2022
Db	1906	CTGCTGGCTTCCCTCGACAGCGTTCGAAAAACAGCGGCGAGTGAAGGCGACACGCTGCC	1965
QY	2023	TATTATGTCGTCGGCGCAATGCGGCACGGACTGTCGCGACGCGGACATTCGCGGCC	2082
Db	1966	TATTATGTCGTCGGCGCAATGCGGCACGGACTGTCGCGACGCGGACATTCGCGGCC	2025
QY	2083	GCCGCTCTGAACACGCCGTAGAACAGGCGGCACCAATCTGGAACCTGATGTCGAA	2142
Db	2026	GCCGCTCTGAACACGCCGTAGAACAGGCGGCACCAATCTGGAACCTGATGTCGAA	2085
QY	2143	CTGGATGCTCCTCGAATCATCGCAACACCCGAGACGGTTGAAACTGCGCGCGCACCGC	2202
Db	2086	CTGGATGCTCCTCGAATCATCGCAACACCCGAGACGGTTGAAACTGCGCGCACCGC	2145
QY	2203	ACAGATATGCGGGGATCCGCCCTACGGCGCAACTTCCGGGACGCGGCGCTACAG	2262
Db	2146	ACAGATATGCGGGGATCCGCCCTACGGCGCAACTTCCGGGACGCGGCGCTACAG	2205
QY	2263	CATGGAATGCGCGCGCGCTGTCAGCATCTTCAACAGTCTCGCGCGTACCGTCTATGCC	2322
Db	2206	CATGGAATGCGCGCGCGCTGTCAGCATCTTCAACAGTCTCGCGCGTACCGTCTATGCC	2265
QY	2323	GACGTACCGCGCGCCATGCCGATATGAGGACGCCCGCTGAAAGCCGTATCGGACGG	2382
Db	2266	GACGTACCGCGCGCCATGCCGATATGAGGACGCCCGCTGAAAGCCGTATCGGACGG	2325
QY	2383	TTGGACCAACAGCTACGGGTCTGCGGTCTATCGCGCAAAACCCAAACAGACGGTGAAG	2442
Db	2326	TTGGACCAACAGCTACGGGTCTGCGGTCTATCGCGCAAAACCCAAACAGACGGTGAAG	2385
QY	2443	TGGGAACAGGCGGTGTTCAAGGCAAAATGCGCGCAGTACCCAAACCGTGGCATTCGC	2502
Db	2386	TGGGAACAGGCGGTGTTCAAGGCAAAATGCGCGCAGTACCCAAACCGTGGCATTCGC	2445
QY	2503	GCGAAACCGCGGAAATACGACAGCAGCGCCACACTGGGCATGGGACACACACATGG	2562
Db	2446	GCGAAACCGCGGAAATACGACAGCAGCGCCACACTGGGCATGGGACACACATGG	2505
QY	2563	AGCGAAACAGTGAATGCAAAACCGACAGCATTAGTCTGTTTCGAGGCATACGGAC	2622
Db	2506	AGCGAAACAGTGAATGCAAAACCGACAGCATTAGTCTGTTTCGAGGCATACGGAC	2565
QY	2623	GATGCGGCGATATCGGTCTCTCAAGCCCTGTTCTCTACGAGCGTACAAAAACAGC	2682
Db	2566	GATGCGGCGATATCGGTCTCTCAAGCCCTGTTCTCTACGAGCGTACAAAAACAGC	2625
QY	2683	ATCAGCCGACAGCAGCGTTCGGGACGAACATGCGGAAGCAGCGTCAACGGCACGCTGATG	2742
Db	2626	ATCAGCCGACAGCAGCGTTCGGGACGAACATGCGGAAGCAGCGTCAACGGCACGCTGATG	2685
QY	2743	CAGCTGGGCGACTGGCGGTGTCACGCTTCGTTGCGCAACGGGAGATTGACGGTC	2802
Db	2686	CAGCTGGGCGACTGGCGGTGTCACGCTTCGTTGCGCAACGGGAGATTGACGGTC	2745
QY	2803	GAAGCGGTCTCGCTACGACCTGCTCAACAGGATGATTCGCCGAAAAAGCGAGTGT	2862
Db	2746	GAAGCGGTCTCGCTACGACCTGCTCAACAGGATGATTCGCCGAAAAAGCGAGTGT	2805
QY	2863	TTGGGCTGAGCGGCAACAGCCCTCACTGAAGGCACACTGGTTCGGACTCGCGGGTCTGAAG	2922
Db	2806	TTGGGCTGAGCGGCAACAGCCCTCACTGAAGGCACACTGGTTCGGACTCGCGGGTCTGAAG	2865
QY	2923	CTGTGCGAACCTTCTAGGATTAACCGTCTCTGTTGCAACGGCGGGGTGGAACGCGAC	2982
Db	2866	CTGTGCGAACCTTCTAGGATTAACCGTCTCTGTTGCAACGGCGGGGTGGAACGCGAC	2925
QY	2983	CTGAACGAGCGGACTACACGTTAAGCGCGGCTTTTACCGGCGGACTGCACGAACCGGC	3042
Db	2926	CTGAACGAGCGGACTACACGTTAAGCGCGGCTTTTACCGGCGGACTGCACGAACCGGC	2985

QY	3043	AAGACGGGGCAGCAATATGCCGACACCGCGCTGTTGCCGGTGTGGCGCGGATGTC	3102
Db	2986	AAGACGGGGCAGCAATATGCCGACACCGCGTCTGTTGCCGGCTGGCGGGATGTC	3045
QY	3103	GAATTCGGCAACGGCTGGAAACGGCTTGGCAGCTTACAGCTACGCCGTTCCAAACAGTAC	3162
Db	3046	GAATTCGGCAACGGCTTGGAAACGGCTTGGCAGCTTACAGCTACGCCGTTCCAAACAGTAC	3105
QY	3163	GGCAACACACGGGAGCAGTGGCGGTAGGCTACCGGTTTC	3201
Db	3106	GGCAACACACGGGAGCAGTGGCGGTAGGCTACCGGTTTC	3144
RESULT 19			
AAS43874			
ID	AAS43874	standard; DNA; 4425 BP.	
XX	AAS43874;		
AC	AAS43874;		
XX			
DT	18-DEC-2001	(first entry)	
XX			
DE	Neisseria meningitidis	fusion protein delta-G983-ORF46.1 DNA.	
XX			
KW	Neisseria gonorrhoeae;	leader peptide; fuslon protein; ORF46.1; ds;	
KW	Neisserial protein.		
XX			
OS	Neisseria meningitidis.		
OS	Synthetic.		
XX			
PN	WO200164922-A2.		
XX			
PD	07-SEP-2001.		
XX			
PF	28-FEB-2001;	2001WO-IB00452.	
XX			
PR	28-FEB-2000;	2000GB-0004695.	
PR	13-NOV-2000;	2000GB-0027675.	
XX			
PA	(CHIR-) CHIRON SPA.		
XX			
PI	Arico MB, Comanducci M, Galeotti C, Masignani V, Guilianì MM;		
PI	Pizza M;		
XX			
DR	WPI; 2001-582163/65.		
DR	P-PSDB; AAU27575.		
XX			
PT	Producing heterologous proteins from Neisseria meningitidis and N.		
PT	gonorrhoeae -		
XX			
PS	Example 15; Page 42-43; 119pp; English.		
XX			
CC	The invention relates to methods for the heterologous expression of		
CC	Neisserial proteins from Neisseria meningitidis and Neisseria		
CC	gonorrhoeae. At least one domain in the protein is deleted, e.g. the		
CC	leader peptide, and may be replaced by a domain from a different		
CC	to make a fusion protein, in order to enhance heterologous expression of		
CC	Neisserial proteins. Also, a region of a protein, such as a poly-glycine		
CC	stretch, can be mutated to enhance expression. The proteins used in the		
CC	processes include ORF46.1, 287, 741, 919, 953, 961 and 983. Sequences		
CC	AAS43868-AAS43905 represent DNA molecules encoding Neisserial proteins		
CC	and peptide regions of proteins of the invention.		
XX			
SQ	Sequence 4425 BP; 1165 A; 1259 C; 1219 G; 782 T; 0 other;		
Query Match 25.9%; Score 831; DB 22; Length 4425;			
Best Local Similarity 99.4%; Pred. No. 0;			
Matches 1231; Conservative 0; Mismatches 8; Indels 0; Gaps 0;			
QY	1963	CTGCTGGCTTCCTCGACAGCGTCGAAAAACAGCGGGCAGTGAAGCGCACACGCTGTCC	2022
Db	1906	CTGCTGGCTTCCTCGACAGCGTCGAAAAACAGCGGGCAGTGAAGCGCACACGCTGTCC	1965
QY	2023	TATTATGTCCTCGCGGCAATGCGGCACGAGTGTCTTCGGCAGCGGCACATTCCGCGGCC	2082

Db	1966	TATTATGTCCTCGCGGCAATGCGGCACGAGTGTCTTCGCGAGCGGCACATTCCGCGGCC	2025
QY	2083	GCCGGTCTGAACACACCGCTAGAACAGAGGGGGCAGCAATCTGAAAAACCTGATGTGCGAA	2142
Db	2026	GCCGGTCTGAACACACCGCTAGAACAGAGGGGGCAGCAATCTGAAAAACCTGATGTGCGAA	2085
QY	2143	CTGGATGCCCTCCGAATCATCCGCAACACCGAGAGCGTTGAAACTCGGGCCGCGACCGC	2202
Db	2086	CTGGATGCCCTCCGAATCATCCGCAACACCGAGAGCGTTGAAACTCGGGCAGCGACCGC	2145
QY	2203	ACAGATATGCCGGGCATCCGCCCTACGGCGCAACTTTTCGCGCAGCGGCGAGCCGTACAG	2262
Db	2146	ACAGATATGCCGGGCATCCGCCCTACGGCGCAACTTTTCGCGCAGCGGCGAGCCGTACAG	2205
QY	2263	CATGCCAATGCCGCCCGACGGGTACGCATCTTCAACAGTCTCGCCCGTACCGTCTATGCC	2322
Db	2206	CATGCCAATGCCGCCCGACGGGTACGCATCTTCAACAGTCTCGCCCGTACCGTCTATGCC	2265
QY	2323	GACAGTACCGCCGCCCATGCCGATATCGAGGAGCGCGGCTGAAAGCCGTATCGGACGGG	2382
Db	2266	GACAGTACCGCCGCCCATGCCGATATCGAGGAGCGCGGCTGAAAGCCGTATCGGACGGG	2325
QY	2383	TTGGACCACAACGCTACGGGTCTTCGCGTCATCGCGCAACCAACAGGACGCTGGAACG	2442
Db	2326	TTGGACCACAACGCGCGGTCTCGCGTCATCGCGCAACCAACAGGACGCTGGAACG	2385
QY	2443	TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCGAGTACCCAAACCGTCGGCATGGC	2502
Db	2386	TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCGAGTACCCAAACCGTCGGCATGGC	2445
QY	2503	GCGAAAAACCGCGAAAAATACGACAGCAGCGCCACACTGGGCGATGGACACACACATGG	2562
Db	2446	GCGAAAAACCGCGAAAAATACGACAGCAGCGCCACACTGGGCGATGGACACACACATGG	2505
QY	2563	AGCGAAAAACAGTGCATAATGCAAAACCGACAGCATTAGTCTGTTTCAGGCGATACGGCAC	2622
Db	2506	AGCGAAAAACAGTGCATAATGCAAAACCGACAGCATTAGTCTGTTTCAGGCGATACGGCAC	2565
QY	2623	GATCGGGGGGATATCGGCTATCTCAAAAGCGCTTCTCTCTACGGAGCGCTACAAAAACAGC	2682
Db	2566	GATCGGGGGGATATCGGCTATCTCAAAAGCGCTTCTCTCTACGGAGCGCTACAAAAACAGC	2625
QY	2683	ATCAGCCGCGACGACCGGTGCGGAGCAACATGCGGAAGGAGCGGTCAACGCGACGCTGATG	2742
Db	2626	ATCAGCCGCGACGACCGGTGCGGAGCAACATGCGGAAGGAGCGGTCAACGCGACGCTGATG	2685
QY	2743	CAGCTGGGGCGACTGGGGCGGTGTCAACGTTCCGTTTTCGCGCAACCGGAGATTTGACGGTC	2802
Db	2686	CAGCTGGGGCGACTGGGGCGGTGTCAACGTTTCCGTTTTCGCGCAACCGGAGATTTGACGGTC	2745
QY	2803	GAAGGGCGTCTGGGCTACGACCTGTCAACAGGATGCAATTCGCCCGAAAAAGCAGTGTCT	2862
Db	2746	GAAGGGCGTCTGGGCTACGACCTGTCTCAACAGGATGCAATTCGCCCGAAAAAGCAGTGTCT	2805
QY	2863	TTGGGCTGGAGGGGCAACAGCCCTCACTGAAGGCACACTGTCGCGACTTCGCGGCTCTGAAG	2922
Db	2806	TTGGGCTGGAGGGGCAACAGCCCTCACTGAAGGCACACTGTCGCGACTTCGCGGCTCTGAAG	2865
QY	2923	CTGTCGCAACCCCTTGAGCGATAAAGCCGCTCCTGTTTTCGCAACGGCGGCGGTGGAAACGCGAC	2982
Db	2866	CTGTCGCAACCCCTTGAGCGATAAAGCCGCTCCTGTTTTCGCAACGGCGGCGGTGGAAACGCGAC	2925
QY	2983	CTGAACGGAGGGGACTACAGGGTAACGGGGGCTTTACCGGGCGCGACTGCGAGCAACCGGC	3042
Db	2926	CTGAACGGAGGGGACTACAGGGTAACGGGGGCTTTACCGGGCGCGACTGCGAGCAACCGGC	2985
QY	3043	AAGACGGGCGCAGCAATATCGCGCACACCGCGCTGGTTGCGCGGTCTGGCGCGGATGTC	3102
Db	2986	AAGACGGGCGCAGCAATATCGCGCACACCGCGCTGGTTGCGCGGTCTGGCGCGGATGTC	3045
QY	3103	GAATTCGGCAACGGCTTGGAAACGGCTTGGGACGTTACAGGTACCGCGGTTCGCAACAGTAC	3162

Db 3046 GAATTCGGCAACGGCTGGAACGGCTTGCGACAGTTACAGCTACGGCGGTTCACAAACAGTAC 3105

Qy 3163 GGCAACACACAGCGGACGAGTCGGCGTAGCTACGGTTTC 3201

Db 3106 GGCAACACACAGCGGACGAGTCGGCGTAGCTACGGTTTC 3144

RESULT 20

AAD17039

ID AAD17039 standard; DNA; 4425 BP.

XX AAD17039;

AC AAD17039;

XX 29-NOV-2001 (first entry)

XX N. meningitidis strain 2996 delta G983-ORF46.1 fusion DNA.

DE Heterologous expression; Neisserial protein; open reading frame; ORF;

XX delta G983-ORF46.1 fusion protein; ds.

KW Neisseria meningitidis 2996.

XX

OS

XX

PH Key Location/Qualifiers

FT CDS 1..4425

FT /\*tag= a

FT /product= "N. meningitidis strain 2996 delta

FT G983-ORF46.1 fusion protein"

XX

PN W0200164920-A2.

XX

XX

PD 07-SEP-2001.

XX

PF 28-FEB-2001; 2001WO-IB00420.

XX

PR 28-FEB-2000; 2000GB-0004695.

PR 13-NOV-2000; 2000GB-0027675.

XX

XX (CHIR-) CHIRON SPA.

PA

XX Arico MB, Comanducci M, Galeotti C, Masignani V, Giuliani MW;

PI Pizza M;

XX

DR WPI; 2001-557776/62.

DR P-PSDB; AAE10022.

XX

PT Heterologous expression for the expression of two or more Neisserial

PT proteins in fused state

XX

PS Example 3; Page 13-14; 52pp; English.

XX

CC The present invention relates to a method for simultaneous heterologous

CC expression of two or more Neisserial proteins which are in a fused

CC state. The method is useful for simultaneous heterologous expression of

CC two or more Neisserial proteins. A protein that may be unstable or

CC poorly expressed on its own is assisted by adding a suitable hybrid

CC partner and commercial manufacture is simplified-only one expression and

CC purification need to be employed in order to produce two separately-

CC useful proteins. The present sequence is a DNA encoding

CC Neisseria meningitidis (serogroup B, strain 2996) delta G983-ORF46.1

CC (open reading frame) fusion protein.

XX

SQ Sequence 4425 BP; 1165 A; 1259 C; 1219 G; 782 T; 0 other;

Query Match 25.9%; Score 831; DB 22; Length 4425;

Best Local Similarity 99.4%; Pred. No. 0;

Matches 1231; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1963 CTGCTGGCTTCCTCGACAGCGTCGAAAAACAGCGGCGAGTGAAGGCGACACGCTGTCC 2022

Db 1906 CTGCTGGCTTCCTCGACAGCGTCGAAAAACAGCGGCGAGTGAAGGCGACACGCTGTCC 1965

Qy 2023 TATTATGTCCTCGCGGCAATCGCGCAGGACTGCTTCGGCAGCGGCACATTCCCGGCC 2082

Db 1966 TATTATGTCCTCGCGCAATGCGGCACGAGTCTGTTGGCAGCGGCACATTCCGCGGCC 2025

Qy 2083 GCCGCTCTGAAACACGCGGTAGAAACAGGGCGGACAGCAATCTGAAAAACCTGATGTCGAA 2142

Db 2026 GCCGCTCTGAAACACGCGGTAGAAACAGGGCGGACAGCAATCTGAAAAACCTGATGTCGAA 2085

Qy 2143 CTGGATGCTCCGAAATCATCCGCAACACCCAGAGAGGGTTGAAACTGGGCGCGCGACGCGC 2202

Db 2086 CTGGATGCTCCGAAATCATCCGCAACACCCAGAGAGGGTTGAAACTGGGCGGACGCGC 2145

Qy 2203 ACAGATATGCCGGGCATCCGCCCTACGGCGCAACTTTCCCGCGAGCGGACGCGTACAG 2262

Db 2146 ACAGATATGCCGGGCATCCGCCCTACGGCGCAACTTTCCCGCGAGCGGACGCGTACAG 2205

Qy 2263 CATCGAATGCCGGGCATCCGCCCTACGGGTGTACGCATCTTCAACAGTCTCGCGCTTACCTGTC 2322

Db 2206 CATCGAATGCCGGGCATCCGCCCTACGGGTGTACGCATCTTCAACAGTCTCGCGCTTACCTGTC 2265

Qy 2323 GACAGTACCGCGCCCATGCGCGATATGACGAGCGCGCGGTGAAAGCCGTATCGGACGGG 2382

Db 2266 GACAGTACCGCGCCCATGCGCGATATGACGAGCGCGCGCTGAAAGCCGTATCGGACGGG 2325

Qy 2383 TTGGACACACAGCGTACGGGTCTGCGGTCTATCGGCAAAACCCACAGGACGGTGAACG 2442

Db 2326 TTGGACACACAGCGGCGGTCTGCGGTCTATCGGCAAAACCCACAGGACGGTGAACG 2385

Qy 2443 TGGAAACAGGGCGGTGTTGAAGGCAAAATGCGCGCAGTACCCAAACCGTCGGCATTTGCC 2502

Db 2386 TGGAAACAGGGCGGTGTTGAAGGCAAAATGCGCGCAGTACCCAAACCGTCGGCATTTGCC 2445

Qy 2503 GCGAAACCGCGCGAAATATACACAGCAGCGCGCCACACTGCGGCATGGGACACAGCACATGG 2562

Db 2446 GCGAAACCGCGCGAAATATACACAGCAGCGCGCCACACTGCGGCATGGGACACAGCACATGG 2505

Qy 2563 AGCGAAACAGTGCAGTGCAGAAACCCACAGCACTTAGTCTGTTGCGAGGATACGGCAC 2622

Db 2506 AGCGAAACAGTGCAGTGCAGAAACCCACAGCACTTAGTCTGTTGCGAGGATACGGCAC 2565

Qy 2623 GATCGGGCGATATCGGCTATCTCAAAAGCGCTGTTCTCTACGGACGCTACAAAAACAGC 2682

Db 2566 GATCGGGCGATATCGGCTATCTCAAAAGCGCTGTTCTCTACGGACGCTACAAAAACAGC 2625

Qy 2683 ATCAGCCGACACGCGGTGCGGACGAAACATCGGAAGCAGCGTCAACGGCAGCGTGATG 2742

Db 2626 ATCAGCCGACACGCGGTGCGGACGAAACATCGGAAGCAGCGTCAACGGCAGCGTGATG 2685

Qy 2743 CAGCTGGCGCACCTGGGCGGTGTCACCGTTCCTGTTGCGCAACGGGAGATTGACGCTC 2802

Db 2686 CAGCTGGCGCACCTGGGCGGTGTCACCGTTCCTGTTGCGCAACGGGAGATTGACGCTC 2745

Qy 2803 GAAGCGGTCTCGGCTACGACCTGCTCAACAGGATGCAATTCGCCGAAAAAGGCAAGTCT 2862

Db 2746 GAAGCGGTCTCGGCTACGACCTGCTCAACAGGATGCAATTCGCCGAAAAAGGCAAGTCT 2805

Qy 2863 TTGGCTGGAGCGGCAACAGCGTCTACTGAAGCAGTGTGTCGACTCGCGGCTCTGAAG 2922

Db 2806 TTGGCTGGAGCGGCAACAGCGTCTACTGAAGCAGTGTGTCGACTCGCGGCTCTGAAG 2865

Qy 2923 CTGTGCAACACCTTGGAGCGATAAAGCCGTCCTGTTGTCACCGGGCGGTGGAACGCGAC 2982

Db 2866 CTGTGCAACACCTTGGAGCGATAAAGCCGTCCTGTTGTCACCGGGCGGTGGAACGCGAC 2925

Qy 2983 CTGAACGGACGCACTACACGGTAACGGGCGGCTTTACCGGCGGACTGCGACAAACCGGC 3042

Db 2926 CTGAACGGACGCACTACACGGTAACGGGCGGCTTTACCGGCGGACTGCGACAAACCGGC 2985

Qy 3043 AAGACGGGGGACGCAATATGCGCGACACCGCGCTGTTGCGGCTCTGGGCGCGGATGTC 3102

Db 2986 AAGACGGGGGACGCAATATGCGCGACACCGCGCTGTTGCGGCGCTGGGCGCGGATGTC 3045

Qy 3103 GAATTCGGCAACGGCTGGAACGGCTTGGCAGCTTACAGCTACGCGCGGTTCACAAACAGTAC 3162

Db 3046 GAATTCGGCAACGGCTGGAACGGCTTGGCAGCTTACAGCTACGCGCGGTTCACAAACAGTAC 3105

QY 3163 GGCAACACAGCGGACGAGTCGGCGTAGGCTACCGGTTTC 3201  
|||||  
Db 3106 GGCAACACAGCGGACGAGTCGGCGTAGGCTACCGGTTTC 3144

## RESULT 21

AAAD17041  
ID AAD17041 standard; DNA; 3798 BP.

XX  
AC AAD17041;

XX  
DT 29-NOV-2001 (first entry)

XX  
DE N. meningitidis strain 2996 delta G983-961 fusion DNA.

XX  
KW Heterologous expression; Neisserial protein;  
delta G983-961 fusion protein; ds.

XX  
OS Neisseria meningitidis 2996.

XX  
FH Key Location/Qualifiers  
CDS 1..3798

FT /\*tag= a  
FT /product= "N. meningitidis strain 2996 delta  
FT G983-961 fusion protein"

FT /transl\_except= (pos:2104..2109, aa:Xaa)  
FT /note= "Xaa corresponds to residues 702-885 of delta  
FT G983-961 fusion protein; There is apparent deletion  
FT of codons corresponding to residues 702-885 of this  
FT protein, which alters the reading frame"

XX  
PN W0200164920-A2.

XX  
PD 07-SEP-2001.

XX  
PF 28-FEB-2001; 2001WO-IB00420.

XX  
PR 28-FEB-2000; 2000GB-0004695.

XX  
PR 13-NOV-2000; 2000GB-0027675.

XX  
PA (CHIR-) CHIRON SPA.

XX  
PI Arico MB, Comanducci M, Galeotti C, Massignani V, Giuliani MM;  
PI Pizza M;

XX  
DR WPI: 2001-557776/62.

XX  
DR P-PSDB; AAE10024.

XX  
PT Heterologous expression for the expression of two or more Neisserial  
PT proteins in fused state

XX  
PS Example 3; Page 16-17; 52pp; English.

XX  
CC The present invention relates to a method for simultaneous heterologous  
CC expression of two or more Neisserial proteins which are in a fused  
CC state. The method is useful for simultaneous heterologous expression of  
CC two or more Neisserial proteins. A protein that may be unstable or  
CC poorly expressed on its own is assisted by adding a suitable hybrid  
CC partner and commercial manufacture is simplified-only one expression and  
CC purification need to be employed in order to produce two separately-  
CC useful proteins. The present sequence is a DNA encoding  
CC Neisseria meningitidis (serogroup B, strain 2996) delta G983-961  
CC fusion protein.

XX  
SQ Sequence 3798 BP; 1079 A; 1020 C; 1007 G; 692 T; 0 other;

Query Match 12.4%; Score 398; DB 22; Length 3798;

Best Local Similarity 99.6%; Pred. No. 5.8e-183;

Matches 498; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1168 TCGGCACCTATGAAGCAAGCGTCCTGTTTCACCCGTACAAACCGGATTCAAATTGCCGGA 1227  
|||||

Db 1111 TCGGCACCTATGAAGCAAGCGTCCTGTTTCACCCGTACAAACCGGATTCAAATTGCCGGA 1170  
QY 1228 ACATCCTTTTCCGACCCCATCGTAACGGCACGGCGCTCTGCTGTCGAGAATACCCG 1287  
|||||  
Db 1171 ACATCCTTTTCCGACCCCATCGTAACGGCACGGCGCTCTGCTGTCGAGAATACCCG 1230  
QY 1288 TGGATGAGCAACGACAACTTGCCTAGCACGCTGCTGACAAACGGCTCAGGACATCGGTGCA 1347  
|||||  
Db 1231 TGGATGAGCAACGACAACTTGCCTAGCACGCTGCTGACGACGGCTCAGGACATCGGTGCA 1290  
QY 1348 GTCGGCGTGGACAGCAAGTTTCGGCTGGGGACTGCTGGATCGGGGTAGGCCATGAACGGA 1407  
|||||  
Db 1291 GTCGGCGTGGACAGCAAGTTTCGGCTGGGGACTGCTGGATCGGGGTAGGCCATGAACGGA 1350  
QY 1408 CCCGCGTCTCTTCCGTTTCGGCGACTTTACGGCGATACGAAAGGTACATCCGATATTGCC 1467  
|||||  
Db 1351 CCCGCGTCTCTTCCGTTTCGGCGACTTTACGGCGATACGAAAGGTACATCCGATATTGCC 1410  
QY 1468 TACTCCTTCCGTAAACGACATTTTCAGGCACGGCGGCTGTATCAAAAAGCGCAGCCAA 1527  
|||||  
Db 1411 TACTCCTTCCGTAAACGACATTTTCAGGCACGGCGGCTGTATCAAAAAGCGCAGCCAA 1470  
QY 1528 CTGCAACTGCACGGCAACAAACACCTATACGGGCAAAACCATTTATCGAAGCGGTTTCGCTG 1587  
Db 1471 CTGCAACTGCACGGCAACAAACACCTATACGGGCAAAACCATTTATCGAAGCGGTTTCGCTG 1530  
QY 1588 GTGTTGTACGGCAACAAACAAATCGGATATCGCGCTCGAACCACCAAGGTGCGCTGATTTAT 1647  
Db 1531 GTGTTGTACGGCAACAAACAAATCGGATATCGCGCTCGAACCACCAAGGTGCGCTGATTTAT 1590  
QY 1648 AACGGGGCGGCATCCCGCGG 1667  
|||||  
Db 1591 AACGGGGCGGCATCCCGCGG 1610

## RESULT 22

AAZ53298

ID AAZ53298 standard; DNA; 526 BP.

XX  
AC AAZ53298;

XX  
DT 21-MAR-2000 (first entry)

XX  
DE Neisseria meningitidis ORF 139 partial DNA sequence SEQ ID NO:545.

XX  
KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;  
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;  
KW antibacterial; gene therapy; ds.

XX  
OS Neisseria meningitidis.

XX  
PN W09957280-A2.

XX  
PD 11-NOV-1999.

XX  
PF 30-APR-1999; 99WO-US09346.

XX  
PR 01-MAY-1998; 98US-0083758.

PR 31-JUL-1998; 98US-0094869.

PR 02-SEP-1998; 98US-0098994.

PR 02-SEP-1998; 98US-0099062.

PR 09-OCT-1998; 98US-0103749.

PR 09-OCT-1998; 98US-0103794.

PR 09-OCT-1998; 98US-0103796.

PR 25-FEB-1999; 99US-0121528.

XX  
(CHIR ) CHIRON CORP.

PA (GENO-) INST GENOMIC RES.

XX  
PI Fraser C, Galeotti C, Grandi G, Hickey E, Massignani V, Mora M;

PI Petersen J, Pizza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;

PI Tettelin H, Venter JC;



```

XX WPI: 2000-365622/31.
XX
XX New polypeptide specific for pathogenic Neisseria useful in therapeutic
XX or preventative vaccines and for diagnosis -
XX PT Claim 1; Page 44; 187pp; French.
XX PS
XX The present sequence represents a genomic fragment of Neisseria
CC meningitidis. The specification describes proteins that are specific
CC for pathogenic strains of Neisseria. The polynucleotides, polypeptides,
CC or their antigenic fragments, are used in vaccines to treat or protect
CC against Neisseria infections, particularly by N. meningitidis. The
CC polynucleotide sequences are also used for recombinant production of
CC the polypeptide and to produce attenuated Neisseria strains that
CC overexpress it, or express it in a non-toxic mutant form.
CC note: the present sequence is included in Claim 1, but it is those
CC sequences that do not include the present sequence that are actually
CC claimed.
XX
XX SQ Sequence 286 BP; 88 A; 73 C; 63 G; 62 T; 0 other;
Query Match 5.5%; Score 177; DB 21; Length 286;
Best Local Similarity 100.0%; Pred. No. 1.6e-75;
Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 865 AATTTCGGAGGAGCAGTAGTACCGCCAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGAC 924
DB |||||||
DB 1 AATTTCGGAGGAGCAGTAGTACCGCCAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGAC 60
QY 925 GAGGGTATCCGCCCTGATGCACAAGAGCGATTACGGCAACTTGCTCTACCACATCCGTAAT 984
DB |||||||
DB 61 GAGGGTATCCGCCCTGATGCACAAGAGCGATTACGGCAACTTGCTCTACCACATCCGTAAT 120
QY 985 AAAAACATGCTTTTCATTTTTTCGGCAAGCAATGACGCACAAGCTCACCCCAACACA 1041
DB |||||||
DB 121 AAAAACATGCTTTTCATTTTTTCGGCAAGCAATGACGCACAAGCTCACCCCAACACA 177
RESULT 25
AAZ53297
ID AAZ53297 standard; DNA; 528 BP.
XX
XX AAZ53297;
XX
XX 21-MAR-2000 (first entry)
XX
XX Neisseria meningitidis ORF 138 partial DNA sequence SEQ ID NO:543.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW antibacterial; gene therapy; ds.
XX
OS Neisseria meningitidis.
XX
XX WO9957280-A2.
XX
XX 11-NOV-1999.
XX
XX 30-APR-1999; 99WO-US09346.
XX
XX 01-MAY-1998; 98US-0083758.
PR 31-JUL-1998; 98US-0094869.
PR 02-SEP-1998; 98US-0098994.
PR 02-SEP-1998; 98US-0099062.
PR 09-OCT-1998; 98US-0103749.
PR 09-OCT-1998; 98US-0103794.
PR 09-OCT-1998; 98US-0103796.
PR 25-FEB-1999; 99US-0121528.
XX
XX (CHIR ) CHIRON CORP.
PA PA
PA (GENO-) INST GENOMIC RES.
XX XX

```

```
PR 09-OCT-1998; 980S-0103796.
PR 25-FEB-1999; 990S-0121528.
XX (CHIR ) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX
PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
PI Petersen J, Pizzo M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
PI Tettelin H, Venter JC;
XX
XX WPI: 2000-062150/05.
DR P-PSDB; AAY74534.
XX
XX Novel Neisserial polypeptides predicted to be useful antigens for
PT vaccines and diagnostics
XX
PS Claim 7; Page 392; 1453pp; English.
XX
CC AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941
CC represent novel Neisseria meningitis and N. gonorrhoea polynucleotides
CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ5473 represent
CC PCR primers used in the exemplification of the present invention. The
CC polypeptides, the polynucleotides, antibodies and compositions of
CC the invention can be used as vaccines, as diagnostic reagents, and as
CC immunogenic compositions. The polypeptides can be used in the
CC manufacture of medicaments for treating or preventing infection due to
CC Neisseria bacteria (e.g. meningitis and septicaemia), to detect the
CC presence of Neisseria bacteria, or to raise antibodies. They may also
CC be used to screen for agonists or antagonists, which may themselves
CC have use as antibacterial agents. The polynucleotides of the invention
CC may also be used in gene therapy protocols.
XX
SQ Sequence 537 BP; 169 A; 136 C; 130 G; 102 T; 0 other;

Query Match 4.7%; Score 149; DB 21; Length 537;
Best Local Similarity 100.0%; Pred. No. 6.6e-62;
Matches 149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 344 TGATCAACCTCAACCTGCAATTGAAGCAGGCTATACAGGCGCGGGTAGAGGTAGGTA 403
Db |
344 TGATCAACCTCAACCTGCAATTGAAGCAGGCTATACAGGCGCGGGTAGAGGTAGGTA 403
QY 404 TCGTCGATACAGCGGAATCCGTCGGCAGCATATCCTTTCCCGAAGTGTATGCGAGAAAAG 463
Db |
404 TCGTCGATACAGCGGAATCCGTCGGCAGCATATCCTTTCCCGAAGTGTATGCGAGAAAAG 463
QY 464 AACACGGCTATACGAAATTAACAAAAC 492
Db |
464 AACACGGCTATACGAAATTAACAAAAC 492

RESULT 27
AAZ53299
ID AAZ53299 standard; DNA; 1365 BP.
XX
AC AAZ53299;
XX
DT 21-MAR-2000 (first entry)
XX
DE Neisseria gonorrhoea ORF 140 partial DNA sequence SEQ ID NO:547.
XX
KW Neisseria meningitidis; Neisseria gonorrhoea; antigen; vaccine;
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW antibacterial; gene therapy; ds.
XX
OS Neisseria gonorrhoea.
XX
PN WO9957280-A2.
XX
PD 11-NOV-1999.
XX
PF 30-APR-1999; 99WO-US09346.
XX
```

```
PR 01-MAY-1998; 980S-0083758.
PR 31-JUL-1998; 980S-0094869.
PR 02-SEP-1998; 980S-0098994.
PR 02-SEP-1998; 980S-0099062.
PR 09-OCT-1998; 980S-0103749.
PR 09-OCT-1998; 980S-0103794.
PR 09-OCT-1998; 980S-0103796.
PR 25-FEB-1999; 990S-0121528.
XX (CHIR ) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX
PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
PI Petersen J, Pizzo M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
PI Tettelin H, Venter JC;
XX
XX WPI: 2000-062150/05.
DR P-PSDB; AAY74537.
XX
XX Novel Neisserial polypeptides predicted to be useful antigens for
PT vaccines and diagnostics
XX
PS Claim 7; Page 394; 1453pp; English.
XX
CC AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941
CC represent novel Neisseria meningitis and N. gonorrhoea polynucleotides
CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ5473 represent
CC PCR primers used in the exemplification of the present invention. The
CC polypeptides, the polynucleotides, antibodies and compositions of
CC the invention can be used as vaccines, as diagnostic reagents, and as
CC immunogenic compositions. The polypeptides can be used in the
CC manufacture of medicaments for treating or preventing infection due to
CC Neisseria bacteria (e.g. meningitis and septicaemia), to detect the
CC presence of Neisseria bacteria, or to raise antibodies. They may also
CC be used to screen for agonists or antagonists, which may themselves
CC have use as antibacterial agents. The polynucleotides of the invention
CC may also be used in gene therapy protocols.
XX
SQ Sequence 1365 BP; 324 A; 405 C; 416 G; 220 T; 0 other;

Query Match 4.1%; Score 131; DB 21; Length 1365;
Best Local Similarity 99.5%; Pred. No. 3.6e-53;
Matches 181; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1960 GGCTCTGCTGGCTTCCTCGACAGCGTCGAAAAACAGCGGCGAGTGAAGCGGCACACGCTG 2019
Db |
121 GGCTCTGCTGGCTTCCTCGACAGCGTCGAAAAACAGCGGCGAGTGAAGCGGCACACGCG 180
QY 2020 TCCTATTATGTCGTCGGGCAATCGGCACGCGACTGCTTCGGCAGCGGCACATTCGCGG 2079
Db |
181 TCCTATTATGTCGTCGGGCAATCGGCACGCGACTGCTTCGGCAGCGGCACATTCGCGG 240
QY 2080 CCCGCCGGTCTGAACACGCCCTAGAACAGGCGCGCAATCTGGAACCTGATGGTC 2139
Db |
241 CCCGCCGGTCTGAACACGCCCTAGAACAGGCGCGCAATCTGGAACCTGATGGTC 300
QY 2140 GA 2141
Db ||
301 GA 302

RESULT 28
AAA81823/c
ID AAA81823 standard; DNA; 476 BP.
XX
AC AAA81823;
XX
DT 04-DEC-2000 (first entry)
XX
DE N. meningitidis partial DNA sequence gnm_370 SEQ ID NO:370.
XX
KW Neisseria meningitidis; Neisseria gonorrhoea; genome; immunogenic;
KW antigen; vaccine; diagnosis; infection; antibacterial; identification;
```



KW	Meningococcus B; MenB; ds.
XX	
OS	Neisseria meningitidis.
XX	
PN	WO200022430-A2.
XX	
PD	20-APR-2000.
XX	
XX	08-OCT-1999; 99WO-US23573.
PF	
XX	09-OCT-1998; 98US-0103794.
PR	
XX	30-APR-1999; 99US-0132068.
PA	(CHIR ) CHIRON CORP.
PI	Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;
PI	Magagnoli V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
PI	Rappuoli R, Pizza M;
XX	
DR	WPI; 2000-318079/27.
XX	
PT	Isolated nucleotide sequences of Neisseria meningitidis which can be
PT	used in the diagnosis and treatment of N. meningitidis infection and
PT	other Neisserial infections, for example, N.gonorrhoea .
XX	
PS	Claim 7; Page 1608; 1760pp; English.
XX	
CC	The present invention describes methods of obtaining immunogenic
CC	proteins from Neisseria genomic sequences. AAA81453 to AAA82414
CC	represent specifically claimed Neisseria meningitidis genomic DNA
CC	sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
CC	Neisseria DNA sequences and their corresponding proteins; AAA81254 to
CC	AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the
CC	isolation of Neisseria meningitidis DNA sequences; and AAA81322 to
CC	AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF
CC	sequences, which are all used in the exemplification of the present
CC	invention. The nucleic acid sequences, protein sequences, and antibodies
CC	against them, can be used in the manufacture of a composition. The
CC	composition can be used as a medicament (or in the manufacture of a
CC	medicament) for treating, preventing or diagnosing infection due to
CC	Neisserial bacteria. For example, some of the identified proteins could
CC	be components of vaccines against Meningococcus B; against all serotypes;
CC	and/or against all pathogenic Neissariae. Identification of sequences
CC	from the bacterium will also facilitate production of biological probes,
CC	particularly organism-specific probes. Attempts to make efficacious
CC	Meningococcus B vaccines have failed mainly due to antigen tolerance.
CC	Multivalent vaccines have also been tried but none have successfully
CC	overcome antigenic variability. The provision of further, complete
CC	sequences may provide an opportunity to identify secreted or surface
CC	exposed proteins that may be presumed targets for the immune system and
CC	which are not antigenically variable or at least more conserved than
CC	other more variable regions.
XX	
SQ	Sequence 476 BP; 84 A; 154 C; 125 G; 111 T; 2 other;
	Query Match 1.8%; Score 57; DB 21; Length 476;
	Best Local Similarity 100.0%; Pred. No. 3.5e-17;
	Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps
Qy	1072 GCTCAAAAGGCAATTATCAGTCGCAGCGGGTAGACCGCAGTGGAAGAAGTTCAAT 1128       
Dd	467 GCTCAAAAGGCAATTATCAGTCGCAGCGGGTAGACCGCAGTGGAAGAAGTTCAAT 411       
RESULT 29	
AAAB2003	
ID	AAAB2003 standard; DNA; 577 BP.
XX	
AC	AAAB2003;
XX	
DT	04-DEC-2000 (first entry)
XX	
DE	N. meningitidis partial DNA sequence qnm_550 SEQ ID NO:550

DT 18-DEC-2001 (first entry)  
XX Neisseria meningitidis B MC58 genomic DNA sequence PCR primer #343.  
DE  
XX  
XX Neisseria gonorrhoeae; leader peptide; fusion protein; ORF46.1; ss;  
KW Neisserial protein; PCR primer.  
XX  
XX Neisseria meningitidis.  
OS Synthetic.  
OS  
XX WO200164922-A2.  
PN  
XX  
XX 07-SEP-2001.  
PD  
XX  
XX 28-FEB-2001; 2001WO-IB00452.  
PF  
XX  
XX 28-FEB-2000; 2000GB-0004695.  
PR  
XX 13-NOV-2000; 2000GB-0027675.  
PR  
XX (CHIR-) CHIRON SPA.  
PA  
XX Arico MB, Comanducci M, Galeotti C, Massignani V, Guillianini MM;  
PI Pizsa M;  
PI  
XX WPI; 2001-582163/65.  
DR  
XX Producing heterologous proteins from Neisseria meningitidis and N.  
PT gonorrhoeae -  
PT  
XX Disclosure; Page 89; 119pp; English.  
PS  
XX The invention relates to methods for the heterologous expression of  
CC Neisserial proteins from Neisseria meningitidis and Neisseria  
CC gonorrhoeae. At least one domain in the protein is deleted, e.g. the  
CC leader peptide, and may be replaced by a domain from a different protein  
CC to make a fusion protein, in order to enhance heterologous expression of  
CC Neisserial proteins. Also, a region of a protein, such as a poly-glycine  
CC stretch, can be mutated to enhance expression. The proteins used in the  
CC processes include ORF46.1, 287, 741, 919, 953, 961 and 983. Sequences  
CC AAS43807-AAS43867 and AAS43906-AAS44358 represent PCR primers used in the  
CC methods of the invention.  
XX  
XX Sequence 50 BP; 15 A; 20 C; 7 G; 8 T; 0 other;  
SQ  
Query Match 1.1%; Score 35; DB 22; Length 50;  
Best Local Similarity 100.0%; Pred. No. 1.8e-06;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 4 CGACGACCCCAACCTCCCTACAAAACTTTCAA 38  
DB 16 CGACGACCCCAACCTCCCTACAAAACTTTCAA 50  
RESULT 31  
ABL05527  
ID ABL05527 standard; cDNA; 2280 BP.  
XX  
AC ABL05527;  
XX  
XX 26-MAR-2002 (first entry)  
DT  
XX  
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 11063.  
DE  
XX  
XX Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical; gene; ss.  
KW  
XX Drosophila melanogaster.  
OS  
XX WO200171042-A2.  
PN  
XX  
XX 27-SEP-2001.  
PD  
XX  
XX 23-MAR-2001; 2001WO-US09231.  
PR  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions -

XX 23-MAR-2000; 2000US-191637P.  
PR 11-JUL-2000; 2000US-0614150.  
XX  
XX (PEKE ) PE CORP NY.  
PA  
XX Venter JC, Adams M, Li PWD, Myers EW;  
PI  
XX WPI; 2001-656860/75.  
DR P-PSDB; ABB61424.  
DR  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions -  
XX  
XX Claim 1; SEQ ID NO 11063; 2lpp + Sequence Listing; English.  
PS  
XX The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABB57737-ABB72072).  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX Sequence 2280 BP; 518 A; 605 C; 561 G; 596 T; 0 other;  
SQ  
Query Match 0.8%; Score 25; DB 23; Length 2280;  
Best Local Similarity 100.0%; Pred. No. 0.12;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 143 GCAGCAACAGCAGAGCAACAACAGC 167  
DB 17 GCAGCAACAGCAGAGCAACAACAGC 41  
RESULT 32  
ABL28074/c  
ID ABL28074 standard; DNA; 3302 BP.  
XX  
XX ABL28074;  
AC  
XX  
XX 26-MAR-2002 (first entry)  
DT  
XX  
XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 35695.  
DE  
XX  
XX Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical; gene; ds.  
KW  
XX Drosophila melanogaster.  
OS  
XX WO200171042-A2.  
PN  
XX  
XX 27-SEP-2001.  
PD  
XX  
XX 23-MAR-2001; 2001WO-US09231.  
PR  
XX  
XX 23-MAR-2000; 2000US-191637P.  
PR 11-JUL-2000; 2000US-0614150.  
XX  
XX (PEKE ) PE CORP NY.  
PA  
XX Venter JC, Adams M, Li PWD, Myers EW;  
PI  
XX WPI; 2001-656860/75.  
DR  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions -

XX PS Claim 1; SEQ ID NO 35695; 21pp + Sequence Listing; English.

XX CC The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from Drosophila. The invention is

CC useful in developmental biology and in elucidating cell signalling and

CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

CC sequences (ABL01840-ABL16175) and the encoded proteins

CC (ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 3302 BP; 765 A; 737 C; 817 G; 983 T; 0 other;

Query Match 0.8%; Score 25; DB 23; Length 3302;

Best Local Similarity 100.0%; Pred. No. 0.12;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 143 GCAGCAACAGCAGCAGCAACACAGC 167

|||||

Db 231 GCAGCAACAGCAGCAGCAACACAGC 207

RESULT 33

ABL05526/c

ID ABL05526 standard; cDNA; 13327 BP.

XX AC ABL05526;

XX DT 26-MAR-2002 (first entry)

XX DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 11060.

XX KW Drosophila; developmental biology; cell signalling; insecticide;

XX KW pharmaceutical; gene; ss.

XX OS Drosophila melanogaster.

XX PN WO200171042-A2.

XX PD 27-SEP-2001.

XX PF 23-MAR-2001; 2001WO-US09231.

XX PR 23-MAR-2000; 2000US-191637P.

XX PR 11-JUL-2000; 2000US-0614150.

XX PA (PEKE ) PE CORP NY.

XX PI Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX DR P-PSDB; ABB61423.

XX PT New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from Drosophila and for elucidating cell signalling and cell-cell

PT interactions -

XX PS Claim 1; SEQ ID NO 11060; 21pp + Sequence Listing; English.

XX CC The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from Drosophila. The invention is

CC useful in developmental biology and in elucidating cell signalling and

CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

CC sequences (ABL01840-ABL16175) and the encoded proteins

CC (ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 13327 BP; 3971 A; 2726 C; 2711 G; 3919 T; 0 other;

Query Match 0.8%; Score 25; DB 23; Length 13327;

Best Local Similarity 100.0%; Pred. No. 0.11;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 143 GCAGCAACAGCAGCAGCAACACAGC 167

|||||

Db 12311 GCAGCAACAGCAGCAGCAACACAGC 12287

RESULT 34

AAA15341/c

ID AAA15341 standard; DNA; 30 BP.

XX AC AAA15341;

XX DT 04-SEP-2000 (first entry)

XX DE PCR primer for a Neisseria pathogenic strain DNA sequence.

XX KW Pathogenic strain; Neisseria; vaccine; Neisseria infection;

XX KW PCR primer; ss.

XX OS Neisseria meningitidis.

XX PN WO200026375-A2.

XX PD 11-MAY-2000.

XX PF 28-OCT-1999; 99WO-FR02643.

XX PR 30-OCT-1998; 98FR-0013693.

XX PA (INRM ) PASTEUR MERIEUX SERUMS & VACCINS SA.

XX PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.

XX PI Aujame L, Bouchardon A, Renauld-Mongenie G, Rokbi B, Nassif X;

XX PI Tinsley C, Perrin A;

XX WPI; 2000-365622/31.

XX PT New polypeptide specific for pathogenic Neisseria useful in therapeutic

PT or preventative vaccines and for diagnosis -

XX PS Example 1; Page 27; 187pp; French.

XX CC The present sequence represents a PCR primer that was used to

CC amplify DNA encoding a protein that is specific for pathogenic

CC strains of Neisseria. The polynucleotides, polypeptides, or their

CC antigenic fragments, are used in vaccines to treat or protect against

CC Neisseria infections, particularly by N. meningitidis. The

CC polynucleotide sequence is also used for recombinant production of

CC the polypeptide and to produce attenuated Neisseria strains that

CC overexpress it, or express it in a non-toxic mutant form.

XX SQ Sequence 30 BP; 7 A; 11 C; 9 G; 3 T; 0 other;

Query Match 0.7%; Score 23; DB 21; Length 30;

Best Local Similarity 100.0%; Pred. No. 1.3;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3181 GTCGCGTAGGTACCGGTTCTG 3203

|||||

Db 30 GTCGCGTAGGTACCGGTTCTG 8

RESULT 35

ABL27930

ID ABL27930 standard; DNA; 2382 BP.

XX

```

AC ABL27930;
XX
XX DT 26-MAR-2002 (first entry)
XX
XX DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 35263.
XX
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX
XX OS Drosophila melanogaster.
XX
XX PN WO200171042-A2.
XX
XX PD 27-SEP-2001.
XX
XX PF 23-MAR-2001; 2001WO-US09231.
XX
XX PR 23-MAR-2000; 2000US-191637P.
XX
XX PR 11-JUL-2000; 2000US-0614150.
XX
XX PA (PEKE ) PE CORP NY.
XX
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX
XX DR WPI; 2001-656860/75.
XX
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX
XX PS Claim 1; SEQ ID NO 35263; 21pp + Sequence Listing; English.
XX
XX CC The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins
XX (ABB57737-ABB72072).
XX
XX CC The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 2382 BP; 685 A; 525 C; 530 G; 642 T; 0 other;

Query Match 0.7%; Score 22; DB 23; Length 2382;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 266 GGGATGCCCAAAATCAATGCCCC 287
Db 2340 GGGATGCCCAAAATCAATGCCCC 2361

RESULT 36
AAAA15340
ID AAA15340 standard; DNA; 34 BP.
XX
XX AC AAA15340;
XX
XX DT 04-SEP-2000 (first entry)
XX
XX DE PCR primer for a Neisseria pathogenic strain DNA sequence.
XX
XX KW Pathogenic strain; Neisseria; vaccine; Neisseria infection;
XX
XX KW PCR primer; ss.
XX
XX OS Neisseria meningitidis.
XX
XX PN WO200026375-A2.
XX
XX DT 11-MAY-2000.

```

```

XX
XX PF 28-OCT-1999; 99WO-FR02643.
XX
XX PR 30-OCT-1998; 98FR-0013693.
XX
XX PA (INMR ) PASTEUR MERIEUX SERUMS & VACCINS SA.
XX
XX PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX
XX PI Aujame L, Bouchardon A, Renauld-Mongenie G, Rokbi B, Nassif X;
XX Tinsley C, Perrin A;
XX
XX DR WPI; 2000-365622/31.
XX
XX PT New polypeptide specific for pathogenic Neisseria useful in therapeutic
XX or preventative vaccines and for diagnosis
XX
XX PS Example 1; Page 27; 187pp; French.
XX
XX CC The present sequence represents a PCR primer that was used to
XX amplify DNA encoding a protein that is specific for pathogenic
XX strains of Neisseria. The polynucleotides, polypeptides, or their
XX antigenic fragments, are used in vaccines to treat or protect against
XX Neisseria infections, particularly by N. meningitidis. The
XX polynucleotide sequence is also used for recombinant production of
XX the polypeptide and to produce attenuated Neisseria strains that
XX overexpress it, or express it in a non-toxic mutant form.
XX
XX SQ Sequence 34 BP; 9 A; 15 C; 5 G; 5 T; 0 other;

Query Match 0.7%; Score 21; DB 21; Length 34;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCGAAGACCCCAACCTTC 21
Db 14 ATGCGAAGACCCCAACCTTC 34

RESULT 37
AAZ55456
ID AAZ55456 standard; DNA; 34 BP.
XX
XX AC AAZ55456;
XX
XX DT 21-MAR-2000 (first entry)
XX
XX DE Neisseria species ORF cloning PCR primer #841.
XX
XX KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
XX
XX KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
XX
XX KW antibacterial; gene therapy; PCR primer; ss.
XX
XX OS Synthetic.
XX
XX OS Neisseria sp.
XX
XX PN WO9957280-A2.
XX
XX DT 11-NOV-1999.
XX
XX PF 30-APR-1999; 99WO-US09346.
XX
XX PR 01-MAY-1998; 98US-0083758.
XX
XX PR 31-JUL-1998; 98US-0094869.
XX
XX PR 02-SEP-1998; 98US-0098994.
XX
XX PR 02-SEP-1998; 98US-0099062.
XX
XX PR 09-OCT-1998; 98US-0103749.
XX
XX PR 09-OCT-1998; 98US-0103794.
XX
XX PR 09-OCT-1998; 98US-0103796.
XX
XX PR 25-FEB-1999; 99US-0121528.
XX
XX (CHIR ) CHIRON CORP.
XX
XX PA (GENO-) INST GENOMIC RES.
XX
XX

```

PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;  
 PI Petersen J, Pizsa M, Rappuoli R, Ratti G, Scalato E, Scarselli M;  
 PI Tettelin H, Venter JC;  
 XX WPI; 2000-062150/05.  
 XX  
 XX Novel Neisserial polypeptides predicted to be useful antigens for  
 PT vaccines and diagnostics -  
 PT  
 XX Example 16; Page 164; 1453pp; English.  
 XX  
 XX AA253015 to AA254536, AA254577 to AA254615, and AA254615 to AA254615  
 CC represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides  
 CC and polypeptides. AA254537 to AA254576 and AA254616 to AA254615 represent  
 CC PCR primers used in the exemplification of the present invention. The  
 CC polypeptides, the polynucleotides, antibodies and compositions of  
 CC the invention can be used as vaccines, as diagnostic reagents, and as  
 CC immunogenic compositions. The polypeptides can be used in the  
 CC manufacture of medicaments for treating or preventing infection due to  
 CC Neisserial bacteria (e.g. meningitis and septicemia), to detect the  
 CC presence of Neisseria bacteria, or to raise antibodies. They may also  
 CC be used to screen for agonists or antagonists, which may themselves  
 CC have use as antibacterial agents. The polynucleotides of the invention  
 CC may also be used in gene therapy protocols.  
 XX  
 XX Sequence 34 BP; 9 A; 10 C; 7 G; 8 T; 0 other;

Query Match 0.6%; Score 20; DB 21; Length 34;  
 Best Local Similarity 100.0%; Pred. No. 36;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 54 GTTAGCTGTTGCCAACAC 73  
 |||||  
 DB 15 GTTAGCTGTTGCCAACAC 34

RESULT 38  
 AAT45009/C  
 ID AAT45009 standard; DNA; 735 BP.  
 XX  
 AC AAT45009;  
 XX  
 XX 13-AUG-1997 (first entry)  
 DT  
 XX Serratia marcescens IAM 13540 acid phosphatase DNA.  
 DE  
 XX IAM 13540; acid phosphatase; production; nucleoside; 5'-phosphate;  
 KW ester; condiment; pharmaceutical; intermediate; ds.  
 XX  
 OS Serratia marcescens.

Key Location/Qualifiers  
 CDS 1..735  
 FT /\*tag= a  
 FT /product= acid\_phosphatase

PN WQ9637603-A1.  
 XX  
 XX 28-NOV-1996.  
 PD  
 XX 24-MAY-1996; 96WO-JP01402.  
 XX  
 XX 26-MAR-1996; 96JP-0094680.  
 PR  
 XX 25-MAY-1995; 95JP-0149781.  
 PR

(AJIN ) AJINOMOTO CO INC.  
 XX  
 XX Asano Y, Mihara Y, Utagawa T, Yamada H;  
 PI  
 XX WPI; 1997-021215/02.  
 DR  
 XX P-PSDB; AAW06460.  
 XX  
 XX Efficient production of nucleoside 5'-phosphate - by reaction of a

PT nucleoside with a phosphoric acid donor in the presence of an acid  
 PT phosphatase  
 XX  
 XX Example 24; Pages 74-75; 95pp; Japanese.  
 XX  
 XX The present sequence encodes the Serratia marcescens IAM 13540  
 CC acid phosphatase (AP), which can be used to produce a nucleoside  
 CC 5'-phosphate ester from the corresponding nucleoside when a  
 CC phosphate donor, e.g. poly-, phenyl- or carbamyl-phosphoric acid,  
 CC is reacted in its presence at pH 3.0 to 5.5. The PA can be used for  
 CC the economic and efficient production of nucleoside 5'-phosphate  
 CC esters for use as condiments, pharmaceuticals and intermediates for  
 CC pharmaceuticals.  
 XX  
 XX Sequence 735 BP; 180 A; 215 C; 198 G; 142 T; 0 other;

Query Match 0.6%; Score 20; DB 18; Length 735;  
 Best Local Similarity 100.0%; Pred. No. 33;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1751 TGCAGCTGGCGGGAAGGT 1770  
 |||||  
 DB 698 TGCAGCTGGCGGGAAGGT 679

RESULT 39  
 AAV43062/C  
 ID AAV43062 standard; DNA; 735 BP.  
 XX  
 AC AAV43062;  
 XX  
 XX 21-OCT-1998 (first entry)  
 DT  
 XX DNA encoding an acid phosphatase enzyme.  
 DE  
 XX Acid phosphatase; preparation; nucleoside 5'-phosphate ester;  
 KW seasoning; intermediate; ds.  
 XX  
 OS Serratia ficaria.

Key Location/Qualifiers  
 CDS 1..735  
 FT /\*tag= a

PN EP857788-A2.  
 XX  
 XX 12-AUG-1998.  
 PD  
 XX 20-NOV-1997; 97EP-0309365.  
 XX  
 XX 18-JUN-1997; 97JP-0161674.  
 PR  
 XX 21-NOV-1996; 96JP-0311103.  
 PR

(AJIN ) AJINOMOTO CO INC.  
 XX  
 XX Asano Y, Mihara Y, Utagawa T, Yamada H;  
 PI  
 XX WPI; 1998-416010/36.  
 DR  
 XX P-PSDB; AAW71033.  
 DR

Preparation of nucleoside 5'-phosphates comprises reacting  
 PT nucleoside with phosphate donor in presence of acid phosphatase -  
 PT used as seasonings or pharmaceutical intermediates  
 XX  
 XX Example 24; Pages 47-48; 83pp; English.

The present sequence encodes an acid phosphatase enzyme. The  
 CC specification describes a method for the preparation of nucleoside  
 CC 5'-phosphate esters. The method comprises reacting a nucleoside with  
 CC a phosphate donor at pH 3.0-5.5 in the presence of an acid phosphatase  
 CC that has been altered to increase its affinity for the nucleoside and/or  
 CC to increase its thermal stability, or in the presence of a microorganism  
 CC that has been transformed with recombinant DNA containing a gene coding

CC for such an acid phosphatase. Nucleoside 5'-phosphates are useful as  
CC seasonings or pharmaceuticals or as intermediates for them.  
CC note: this sequence appears to be claimed (claim 6), but as the claim  
CC refers to amino acid sequences, it is clear that the corresponding  
CC protein is being claimed.  
XX  
SQ Sequence 735 BP; 180 A; 215 C; 198 G; 142 T; 0 other;  
Query Match 0.6%; Score 20; DB 19; Length 735;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1751 TGCAGCTGGCGCGGAAGT 1770  
|||||  
Db 698 TGCAGCTGGCGCGGAAGT 679  
RESULT 40  
ABQ32992/c  
ID ABQ32992 standard; DNA; 1105 BP.  
XX  
AC ABQ32992;  
XX  
DT 12-JUL-2002 (first entry)  
XX  
DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 19583.  
XX  
KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;  
KW drug; side effect; cancer; central nervous system; cardiovascular;  
KW gastrointestinal; respiratory system; single nucleotide polymorphism;  
KW SNP; cell differentiation; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200218632-A2.  
XX  
PD 07-MAR-2002.  
XX  
PF 01-SEP-2001; 2001WO-EP10074.  
XX  
PR 01-SEP-2000; 2000DE-1043826.  
PR 05-SEP-2000; 2000DE-1044543.  
XX  
PA (EPIC-) EPIGENOMICS AG.  
XX  
PI Olek A, Piepenbrock C, Berlin K, Guetig D;  
XX  
WPI: 2002-371829/40.  
DR  
PT Determining the degree of cytosine methylation in genomic DNA, useful  
PT for diagnosis and prognosis, comprises selective hybridization of  
PT amplicons from chemically treated DNA  
XX  
PS Claim 12: 56pp + Sequence Listing; 56pp; German.  
XX  
CC This invention describes a novel method for determining the degree of  
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a  
CC genomic sample of DNA. The sample is treated chemically to convert  
CC cytosine (C) but not methylated C, to uracil, then part of the genomic  
CC DNA that contains the target C is amplified to form a labeled amplicon.  
CC The amplicon is hybridised to two classes, each with at least one  
CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers  
CC and the degree of hybridisation to both classes is determined from the  
CC label on the amplicon. From the ratio of labels hybridised to the two  
CC classes of oligomers, the degree of methylation is calculated. The method  
CC is used: (i) for diagnosis and/or prognosis of side effects of  
CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders  
CC of the central nervous, cardiovascular, gastrointestinal and respiratory  
CC systems etc., particularly by detecting mutations or single nucleotide  
CC polymorphisms (SNP's); and (ii) for differentiation of cell or tissue  
CC types and for investigating cell differentiation. The method allows the  
CC methylation status of many C residues to be determined simultaneously.  
CC ABQ13410-ABQ54121 represent genomic DNA sequences used to illustrate the

CC method for determining the degree of cytosine methylation described in  
CC the disclosure of the invention.  
XX  
SQ Sequence 1105 BP; 144 A; 147 C; 416 G; 398 T; 0 other;  
Query Match 0.6%; Score 20; DB 24; Length 1105;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 475 RACGAAATTTACAAAACTA 494  
|||||  
Db 353 AACGAAATTTACAAAACTA 334  
Search completed: January 27, 2003, 20:57:25  
Job time : 1443 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 27, 2003, 16:40:01 ; Search time 5468 seconds  
(without alignments)  
17052.924 Million cell updates/sec

Title: US-09-830-433a-7  
Perfect score: 3204  
Sequence: 1 atgcgaacgaccccaacctt.....gcgtaggctaccggtctga 3204

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.sts.\*
- 12: gb.sy.\*
- 13: gb.un.\*
- 14: gb.vi.\*
- 15: em.ba.\*
- 16: em.fun.\*
- 17: em.hum.\*
- 18: em.in.\*
- 19: em.mu.\*
- 20: em.om.\*
- 21: em.or.\*
- 22: em.ov.\*
- 23: em.pat.\*
- 24: em.ph.\*
- 25: em.pl.\*
- 26: em.ro.\*
- 27: em.sts.\*
- 28: em.un.\*
- 29: em.vi.\*
- 30: em.htg\_hum.\*
- 31: em.htg\_inv.\*
- 32: em.htg\_other.\*
- 33: em.htg\_mus.\*
- 34: em.htg\_pln.\*
- 35: em.htg\_rod.\*
- 36: em.htg\_mam.\*
- 37: em.htg\_vrt.\*
- 38: em.sv.\*
- 39: em.htgo\_hum.\*
- 40: em.htgo\_mus.\*
- 41: em.htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3204	100.0	3204	6	AX024064	AX024064 Sequence
2	3202.4	100.0	3204	1	AX311654	AX311654 Neisseria
3	3160.6	98.6	349061	1	NME222491	AL162753 Neisseria
4	3062.2	95.6	10591	1	AE002545	AE002545 Neisseria
5	3062.2	95.6	349980	6	AX044034	AX044034 Neisseria
6	3042.4	95.0	3254	1	NME277537	AX277537 Neisseria
7	2986.2	93.2	4218	6	AX236469	AX236469 Sequence
8	2986.2	93.2	4218	6	AX239755	AX239755 Sequence
9	2983.8	93.1	4335	6	AX236457	AX236457 Sequence
10	2983.8	93.1	4335	6	AX239743	AX239743 Sequence
11	2983.2	93.1	3939	6	AX236419	AX236419 Sequence
12	2983.2	93.1	3939	6	AX239729	AX239729 Sequence
13	2983.2	93.1	4170	6	AX236463	AX236463 Sequence
14	2983.2	93.1	4170	6	AX239749	AX239749 Sequence
15	2971.8	92.8	3939	6	AX236409	AX236409 Sequence
16	2971.8	92.8	3939	6	AX239717	AX239717 Sequence
17	2971.8	92.8	4179	6	AX236413	AX236413 Sequence
18	2971.8	92.8	4179	6	AX239721	AX239721 Sequence
19	2971.8	92.8	4344	6	AX236411	AX236411 Sequence
20	2971.8	92.8	4344	6	AX239719	AX239719 Sequence
21	2971.8	92.8	4425	6	AX236407	AX236407 Sequence
22	2971.8	92.8	4425	6	AX239715	AX239715 Sequence
23	576.4	18.0	600	1	AF169448	AF169448 Neisseria
24	270.8	8.5	286	1	AF169473	AF169473 Neisseria
25	270.8	8.5	286	6	A68924	A68924 Sequence 95
26	147.8	4.6	3493	1	BPE318229	AJ318229 Bordetell
27	132.4	4.1	11292	1	AE011850	AE011850 Xanthomon
28	130.2	4.1	11019	1	AE012308	AE012308 Xanthomon
29	81.6	2.5	1434	1	NGOPAK	X52364 N.gonorrhoe
30	77.4	2.4	10807	1	AE011766	AE011766 Xanthomon
31	77	2.4	10029	1	AE012229	AE012229 Xanthomon
32	72.8	2.3	10495	1	AE003880	AE003880 Xylella f
33	72.8	2.3	14988	1	AB015053	AB015053 Pseudomon
34	71.6	2.2	12152	1	AF286062	AF286062 Pseudomon
35	71.2	2.2	11186	1	AE003939	AE003939 Xylella f
36	70.8	2.2	215050	1	AL646057	AL646057 Ralstonia
37	68	2.1	7030	1	SMASSPH	D78380 Serratia ma
38	67	2.1	11476	1	AE004006	AE004006 Xylella f
39	65	2.0	5423	1	AF216702	AF216702 Pseudomon
40	62	1.9	10732	6	E32986	E32986 Gene encodi
41	60.2	1.9	67200	1	MTV017	AL021897 Mycobacte
42	56	1.7	125020	9	AF429315	AF429315 Homo sapi
43	53.2	1.7	125020	9	AF429315	AF429315 Homo sapi
44	53	1.7	3510	1	PT07828	AJ007828 Pseudomon
45	53	1.7	10165	1	AE001935	AE001935 Deinococc

ALIGNMENTS

RESULT 1	AX024064	Sequence 7 from Patent FR2785293.	3204 bp	DNA	linear	PAT 15-SEP-2000
LOCUS	AX024064	Sequence 7 from Patent FR2785293.	3204 bp	DNA	linear	PAT 15-SEP-2000
DEFINITION	AX024064	Sequence 7 from Patent FR2785293.	3204 bp	DNA	linear	PAT 15-SEP-2000
ACCESSION	AX024064	Sequence 7 from Patent FR2785293.	3204 bp	DNA	linear	PAT 15-SEP-2000
VERSION	AX024064.1	GI:10184376	3204 bp	DNA	linear	PAT 15-SEP-2000
KEYWORDS	AX024064.1	GI:10184376	3204 bp	DNA	linear	PAT 15-SEP-2000
SOURCE	AX024064.1	GI:10184376	3204 bp	DNA	linear	PAT 15-SEP-2000
ORGANISM	AX024064.1	GI:10184376	3204 bp	DNA	linear	PAT 15-SEP-2000
REFERENCE	AX024064.1	GI:10184376	3204 bp	DNA	linear	PAT 15-SEP-2000
AUTHORS	AX024064.1	GI:10184376	3204 bp	DNA	linear	PAT 15-SEP-2000
JOURNAL	AX024064.1	GI:10184376	3204 bp	DNA	linear	PAT 15-SEP-2000

FEATURES		PASTEUR MERIEUX SERUMS VACC (FR)	
source	1..3204	Location/Qualifiers	
	/organism="Neisseria meningitidis"		
	/db_xref="taxon:487"		
CDS	1..3204		
	/note="unnamed protein product"		
	/codon_start=1		
	/transl_table=11		
	/protein_id="CAC08908.1"		
	/db_xref="GI:10184377"		
	/translation="MRTPTPTFKTPAAMALAVATLTSACLGGGGTSPDFNAGG		
	TGIGSNRATTAASAYAGIKNEMCKDRSMLCAGRDDVAVTRDAKINAPPNLHT		
	GFTNPNDAYKNILNLPKALTEAGVTGVEVIVDTGESVGSISFPLEYKDHVSHIIG		
	NYKNYAYMEKEPEDGCKDIKASDFDEAVIEPEAKPTDI RHVKETGIDVVSIIIG		
	GRSDGRPAGGIADATLHIMNHDGKTNEIMSAAIRNANVKIGERGVRIYNSFGTT		
	SRAGTADFOIANSEEQYQALAIYSGGDKTDGIRLUMQSDYGNLSYHIRNNMLFI		
	FASNDAAQPNLTLLLPFEYKDAQGIITVAGVDRSEKFNESHCGITAMWLSAP		
	YESVRFRTNPITQIAGTSFSAPIVTFQALLLQKYPWMSNDMLRTTLLTAQDLAG		
	GVDSKFGWLLDAGKAMNGSPAPFGFTADTKGTSIDIAYSFRNDISGTLGKKGGS		
	QLOLHGNTVYTKTILBEGSILVLYGNKSDMRVETKCALIYNGAASGSLNSDGIYVL		
	ADTRSCANETVHIKBDLQGGEGTLYRLCKLLKVDGTAMTGGKLYMSARGKAGYVL		
	NRTGQRPVFLSAKBDLQGGEGTLYRLCKLLKVDGTAMTGGKLYMSARGKAGYVL		
	ARTASAAHAPAGLKHAVEGGSNLENLMAVELDASSESATPETVETFAAADTMDPGI		
	RPYGATPRAAAVQAHANAADGVRFNLSLAATVYADSTAAHADQKRVKAVSDGLDH		
	ATGLURVTAOTQODGTWEQGVKMRGSTQTVGIAAKTGENTAAATLHGMHSTWSE		
	NSANAKTDSISLPAGIRHDAGDILGYLKLFSYGRYKNSISRSTGADHAEHGSYNTLM		
	QLGNALGVNPFPAATGDLTVBEGGLRYDLKQDAFAEKGSALHSGNSLSTEGTLVLAG		
	LKLSQPLSDKVLATAGVERDUNGRTYTVTGFGTATATGKTGARNMPTHLVLVAGL		
	GADVEFGNGWGLFARYAGSKQYGNHSGRVGVGRF"		
BASE COUNT	831 a	909 c	899 g 565 t
ORIGIN			
Query Match	100.0%;	Score 3204;	DB 6; Length 3204;
Best Local Similarity	100.0%;	Pred. No. 0;	
Matches 3204;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
Qy	1	ATGCGAAGCAGCCCAACCTTCCTACAAAACTTTCAAACCGGCTGCCATGGCGTTAGCT	60
Db	1	ATGCGAAGCAGCCCAACCTTCCTACAAAACTTTCAAACCGGCTGCCATGGCGTTAGCT	60
Qy	61	GTTCGCAACACACTTTCGCTGTCTAGGCGCGCGCGGCGGCACTTTCGCGCGGCACTTC	120
Db	61	GTTCGCAACACACTTTCGCTGTCTAGGCGCGCGCGGCGGCACTTTCGCGCGGCACTTC	120
Qy	121	AATGAGCGCGCAGCGGTATCGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	180
Db	121	AATGAGCGCGCAGCGGTATCGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	180
Qy	181	GTATCTTACGCGGTATCAAGAACGAAATGTGCAAGACAGAGCATGCTCTGTGCGCGGT	240
Db	181	GTATCTTACGCGGTATCAAGAACGAAATGTGCAAGACAGAGCATGCTCTGTGCGCGGT	240
Qy	241	CGGGATGAGCTTTCGCGGTATACAGACAGGATGCCAAATCAATTCGCCCCCCCCGAACTG	300
Db	241	CGGGATGAGCTTTCGCGGTATACAGACAGGATGCCAAATCAATTCGCCCCCCCCGAACTG	300
Qy	301	CATACCGGAGCTTTACAAACCCAAATGACGCATACAGAAATTTGATCAACCTCAAACT	360
Db	301	CATACCGGAGCTTTACAAACCCAAATGACGCATACAGAAATTTGATCAACCTCAAACT	360
Qy	361	GCAATTTGAAGCAGGTATACAGGACCGGGGTAGAGTAGGTATGCTGATACAGGCGAA	420
Db	361	GCAATTTGAAGCAGGTATACAGGACCGGGGTAGAGTAGGTATGCTGATACAGGCGAA	420
Qy	421	TCCGTCGCGCAGCATATCCTTTCCGAACTGTATGCGAAGAAACACACGGCTTATAACGAA	480
Db	421	TCCGTCGCGCAGCATATCCTTTCCGAACTGTATGCGAAGAAACACACGGCTTATAACGAA	480
Qy	481	AATTACAAAACACTATACGGCGTATATCGGAGGAGGCGCCTGAAGACGGGCGGTAAA	540
Db	481	AATTACAAAACACTATACGGCGTATATCGGAGGAGGCGCCTGAAGACGGGCGGTAAA	540
Qy	541	GACATTAAGCTTCTTTTCGACGATGAGGCCGCTTATAGAGACTGAAGCAAAAGCCGAGAT	600
Db	541	GACATTAAGCTTCTTTTCGACGATGAGGCCGCTTATAGAGACTGAAGCAAAAGCCGAGAT	600
Qy	601	ATCCGCCACGTAAGAAATCGGACACATCGATGTGCTCCCATATATTTGGCGGGCGGT	660
Db	601	ATCCGCCACGTAAGAAATCGGACACATCGATGTGCTCCCATATATTTGGCGGGCGGT	660
Qy	661	TCCGTGGACGCGACACCTTCGAGCGGTATTTCGCCCGCATGCGCTACACATATGAAT	720
Db	661	TCCGTGGACGCGACACCTTCGAGCGGTATTTCGCCCGCATGCGCTACACATATGAAT	720
Qy	721	ACGCATGATGAACCAAGAAATATATGTCTCAGCCATCCGCAATTCGATGGGTCAAG	780
Db	721	ACGCATGATGAACCAAGAAATATATGTCTCAGCCATCCGCAATTCGATGGGTCAAG	780
Qy	781	CTGGCGGAACCTGGCGGTGCGCATCGTCAATAACAGTTTGTGGAACAACATCGAGGCGAGC	840
Db	781	CTGGCGGAACCTGGCGGTGCGCATCGTCAATAACAGTTTGTGGAACAACATCGAGGCGAGC	840
Qy	841	ACTCCGACCACTTTCCAAATAGCCAATTCGAGGAGCAGTACCGCAAGCGTTCTCGCC	900
Db	841	ACTCCGACCACTTTCCAAATAGCCAATTCGAGGAGCAGTACCGCAAGCGTTCTCGCC	900
Qy	901	TATTCGCGCGGTGATAAACAAGAGAGGATATCGCCCTGATGCAACAGAGGATTCACGC	960
Db	901	TATTCGCGCGGTGATAAACAAGAGAGGATATCGCCCTGATGCAACAGAGGATTCACGC	960
Qy	961	AACTTGTCTTACACATCCGTAATAAACAATGCTTTTTCATTTTTCGGGAAGCAATGAC	1020
Db	961	AACTTGTCTTACACATCCGTAATAAACAATGCTTTTTCATTTTTCGGGAAGCAATGAC	1020
Qy	1021	GCACAGCTCAGCCCAACACACTGACCCCTATTGCTTTTATGAAAAGAGTCTCAAAA	1080
Db	1021	GCACAGCTCAGCCCAACACACTGACCCCTATTGCTTTTATGAAAAGAGTCTCAAAA	1080
Qy	1081	GGCATTATACAGTCGCGAGGCTAGACCGCAGTAGACCGCAGTGAAGAAAAGTTCAATGGCTCCAAACCAT	1140
Db	1081	GGCATTATACAGTCGCGAGGCTAGACCGCAGTAGACCGCAGTGAAGAAAAGTTCAATGGCTCCAAACCAT	1140
Qy	1141	TGCGGAATTTACTGCGATGTGCTGCTATCGGACCCCTATGAGCAAGGCTCGTTTACCC	1200
Db	1141	TGCGGAATTTACTGCGATGTGCTGCTATCGGACCCCTATGAGCAAGGCTCGTTTACCC	1200
Qy	1201	CGTACAAACCCGATTCAAATTTCCGGAACATCTTTTCCGACCCATCGTAACCGGCACG	1260
Db	1201	CGTACAAACCCGATTCAAATTTCCGGAACATCTTTTCCGACCCATCGTAACCGGCACG	1260
Qy	1261	CGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACTCGGTACCCAGCTG	1320
Db	1261	CGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACTCGGTACCCAGCTG	1320
Qy	1321	CTGACACCGCTCAGGACATCGGTGAGTGGGTGGACAGCAAGTTGCGGTGGGAGCTG	1380
Db	1321	CTGACACCGCTCAGGACATCGGTGAGTGGGTGGACAGCAAGTTGCGGTGGGAGCTG	1380
Qy	1381	CTGATGCGGGTAAAGCCATGAAGGACCCGCTCTTTTCCGTCGCGGACTTTTACCCGC	1440
Db	1381	CTGATGCGGGTAAAGCCATGAAGGACCCGCTCTTTTCCGTCGCGGACTTTTACCCGC	1440
Qy	1441	GATACGAAAGGTACATCCGATATTCCTTCTCTACTCTTCCGTAACGACATTTTCAGGACGGGC	1500
Db	1441	GATACGAAAGGTACATCCGATATTCCTTCTCTACTCTTCCGTAACGACATTTTCAGGACGGGC	1500
Qy	1501	GGCCTGATCAAAAAGGCGGAGCCAACTGCACTGACGGCAACCAACCTTATACGGGC	1560
Db	1501	GGCCTGATCAAAAAGGCGGAGCCAACTGCACTGACGGCAACCAACCTTATACGGGC	1560
Qy	1561	AAACCATTTATCGAAGCGGTTGCTGCTGTTGTACGGCAACCAACCTTATACGGGC	1620
Db	1561	AAACCATTTATCGAAGCGGTTGCTGCTGTTGTACGGCAACCAACCTTATACGGGC	1620
Qy	1621	GTGGAACCAAGGTGCGCTGATTTTAAACGGGCGGATCCGGCGGTAGCCTGAACAGC	1680



```
Db 1621 GTGCAAAACCAAGTGCCTGATTATTAACGGGGCGGCATCCGCGGTACGCTGAACAGC 1680
QY 1681 GACGGCATTCCTATCTGGCAGATACCGACCGATCCGGCGCAACAAACCGTGCACATC 1740
Db 1681 GACGGCATTCCTATCTGGCAGATACCGACCGATCCGGCGCAACAAACCGTGCACATC 1740
QY 1741 AAAGGCGATCTCAGCTGGCGCGCAAGGTACGCTGTACACAGCTTTGGGCAAACTGCTG 1800
Db 1741 AAAGGCGATCTCAGCTGGCGCGCAAGGTACGCTGTACACAGCTTTGGGCAAACTGCTG 1800
QY 1801 AAAGTGGACGGTACGGGATACCGCGCGCAAGGTACGCTGTACATGTCCGACCGGCAAGGG 1860
Db 1801 AAAGTGGACGGTACGGGATACCGCGCGCAAGGTACGCTGTACATGTCCGACCGGCAAGGG 1860
QY 1861 GCAGGCTATCTCAACCGTACCGGACACAGTGTTCCTCTCTAGTGCAGCGCAAAATCGGG 1920
Db 1861 GCAGGCTATCTCAACCGTACCGGACACAGTGTTCCTCTCTAGTGCAGCGCAAAATCGGG 1920
QY 1921 CGGGATATTCCTTTCACAAACATCGAAACCGAGCGGTGCTGCTGCTTCCCTCGAC 1980
Db 1921 CGGGATATTCCTTTCACAAACATCGAAACCGAGCGGTGCTGCTGCTTCCCTCGAC 1980
QY 1981 AGCGTGAACAAACAGGGCAGTGAAGCGACACGCTGTCCTATTATGTCCGTGCGGGC 2040
Db 1981 AGCGTGAACAAACAGGGCAGTGAAGCGACACGCTGTCCTATTATGTCCGTGCGGGC 2040
QY 2041 AATGCGGACGGACTGCTTCGGCAGCGGCACATTTCGCGCGCGCGCTCTGAACACGCGC 2100
Db 2041 AATGCGGACGGACTGCTTCGGCAGCGGCACATTTCGCGCGCGCGCTCTGAACACGCGC 2100
QY 2101 GTAGAACAGGGCGGACGAACTCTGAAACCTGATGCTGAACCTGATGCTCCGGAATCA 2160
Db 2101 GTAGAACAGGGCGGACGAACTCTGAAACCTGATGCTGAACCTGATGCTCCGGAATCA 2160
QY 2161 TCGGCAACACCCGAGACGGTTGAACCTGCGGCGCGGACCGACAGATATGCGCGGCATC 2220
Db 2161 TCGGCAACACCCGAGACGGTTGAACCTGCGGCGCGGACCGACAGATATGCGCGGCATC 2220
QY 2221 CGGCCCTACGGCGCAACTTTCCGCGCAGCGGACGCGTACAGCATCGAATGCGCGGCAC 2280
Db 2221 CGGCCCTACGGCGCAACTTTCCGCGCAGCGGACGCGTACAGCATCGAATGCGCGGCAC 2280
QY 2281 GGTGTACGCATCTTCAACAGTCTCGCGCTACCGTCTATGCCGACAGTACCGCGGCCAT 2340
Db 2281 GGTGTACGCATCTTCAACAGTCTCGCGCTACCGTCTATGCCGACAGTACCGCGGCCAT 2340
QY 2341 GCGGATATCGAGGACCGCGGCTGAAGCGGTATCGGACGGGTGGACCAACAGCTACG 2400
Db 2341 GCGGATATCGAGGACCGCGGCTGAAGCGGTATCGGACGGGTGGACCAACAGCTACG 2400
QY 2401 GGTCTCGGCTCATCGCGCAACCCCAACAGGACGGTGGAACTGGGAACAGGCGGTGT 2460
Db 2401 GGTCTCGGCTCATCGCGCAACCCCAACAGGACGGTGGAACTGGGAACAGGCGGTGT 2460
QY 2461 GAAGGCAAAATCGCGGCACTACCAACCGTCCGCAATTTGCCGCAAAACCGCGCAAAAT 2520
Db 2461 GAAGGCAAAATCGCGGCACTACCAACCGTCCGCAATTTGCCGCAAAACCGCGCAAAAT 2520
QY 2521 ACAGACAGACCGCCACACTGGGCATGGGACACAGACATGGAGCGGAAACAGCTGCAAT 2580
Db 2521 ACAGACAGACCGCCACACTGGGCATGGGACACAGACATGGAGCGGAAACAGCTGCAAT 2580
QY 2581 GCAAAAACCCAGACGATATGCTGTTTGCAGGCATACGCGACGATCGGCGGATATCGGC 2640
Db 2581 GCAAAAACCCAGACGATATGCTGTTTGCAGGCATACGCGACGATCGGCGGATATCGGC 2640
QY 2641 TATCTCAAGGGCGTGTCTCTACGACGCTACAAAACAGCATACGCGCAGCACCGGT 2700
Db 2641 TATCTCAAGGGCGTGTCTCTACGACGCTACAAAACAGCATACGCGCAGCACCGGT 2700
QY 2701 GCGGACGACATCGGGAAGCGACGCTCAACGCGACGCTGATGACGCTGGCGCACTGGGC 2760
|||||
```

```
Db 2701 GCGGACGACATCGGAAAGCAGCGTCAACGGCACGCTGATGAGCTGGCGCACTGGC 2760
QY 2761 GGTGTCAACCTTCGTTTCCGCAACGGGAGATTGACGCTCGAAGCGCTCTCGCGTAC 2820
Db 2761 GGTGTCAACCTTCGTTTCCGCAACGGGAGATTGACGCTCGAAGCGCTCTCGCGTAC 2820
QY 2821 GACCTGCTCAACAGGATGATTCGCCGGAAGGAGGAGTCTTTGGGCTGGAGCGCAAC 2880
Db 2821 GACCTGCTCAACAGGATGATTCGCCGGAAGGAGGAGTCTTTGGGCTGGAGCGCAAC 2880
QY 2881 AGCCTCACTGAAGCAGCTGTCGGACTCGGGGCTGAAAGCTGTCGCAACCTTTGAGC 2940
Db 2881 AGCCTCACTGAAGCAGCTGTCGGACTCGGGGCTGAAAGCTGTCGCAACCTTTGAGC 2940
QY 2941 GATAAAGCCCTCTCTGTTTGAACGCGGGGCTGGAACCGACCTGAAACGACGCACTAC 3000
Db 2941 GATAAAGCCCTCTCTGTTTGAACGCGGGGCTGGAACCGACCTGAAACGACGCACTAC 3000
QY 3001 ACGGTAAACGGCGGCTTTACCGCGGACTGACAGCAACCGGCAAGCGGGGACGCAAT 3060
Db 3001 ACGGTAAACGGCGGCTTTACCGCGGACTGACAGCAACCGGCAAGCGGGGACGCAAT 3060
QY 3061 ATCCGCGACACCGCTGTTGCGGCTCGGGGCTGCGGATGTCGAAATTCGCAACGGCTGG 3120
Db 3061 ATCCGCGACACCGCTGTTGCGGCTCGGGGCTGCGGATGTCGAAATTCGCAACGGCTGG 3120
QY 3121 AACGGCTTGGCAGCTTACAGCTACGCGGTTCCAAACAGTACGCAACCGGCAAGCGGAC 3180
Db 3121 AACGGCTTGGCAGCTTACAGCTACGCGGTTCCAAACAGTACGCAACCGGCAAGCGGAC 3180
QY 3181 GTCGGCTAGGCTACCGGTTCTGA 3204
Db 3181 GTCGGCTAGGCTACCGGTTCTGA 3204

RESULT 2
NME311654
LOCUS
DEFINITION NME311654 3204 bp DNA linear BCT 24-JUL-2002
Neisseria meningitidis ausp gene for autotransporter serine
protease.
ACCESSION AJ311654
VERSION AJ311654.1 GI:13397949
KEYWORDS ausp gene; autotransporter serine protease.
SOURCE Neisseria meningitidis.
ORGANISM Neisseria meningitidis
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
Neisseria.
REFERENCE
1
AUTHORS Turner,D.P., Wooldridge,K.G. and Ala'Aldeen,D.A.
TITLE Autotransported Serine Protease A of Neisseria meningitidis: an
Immunogenic, Surface-Exposed Outer Membrane, and Secreted Protein
Infect. Immun. 70 (8), 4447-4461 (2002)
JOURNAL 22112893
MEDLINE 2 (bases 1 to 3204)
AUTHORS Ala'Aldeen,D.A.
TITLE Direct Submission
JOURNAL Submitted (13-MAR-2001) Ala'Aldeen D.A., Microbiology, University
of Nottingham, University Hospital, Nottingham, NG7 2UH, UNITED
KINGDOM
COMMENT Related sequence: AJ277537.
FEATURES
source
1. .3204
/organism="Neisseria meningitidis"
/strain="Group C EF37"
/isolate="24181"
/db_xref="taxon:487"
/country="Mali"
1. .3204
/gene="ausp"
1. .3204
/gene="ausp"
/codon_start=1

gene
CDS
```

/transl\_table=11  
/product="autotransporter serine protease (Aus?)"  
/protein\_id="CAC34605.1"  
/db\_xref="GI:13397950"  
/db\_xref="SPTREMBL:Q9AE78"

/translation="MRITPTPTPTKFFPAAMALAVAVTSLACLGSGGGTSAPDFNAGG  
TGIGNSRATTAKSAAYSAIKNEMCKDRSMCAGRDVAVTDRDAKINAPPNLHT  
GDFNPNDAYKNLNLNPAIEAGYTGKVEVIGVDTGESVGSISFPPELYGRKERGYNE  
NKNYATPMRKEAPDEGGDKIKASFDDEAVEIETEAQPTDIRHVKETIGHIDVNSHIIG  
GRSVDGPAGIADATLHINWHDGKNEIMSAIIRNAWKLGKRGVRIVNNSEFGTT  
SPAGTADHFQIANSEEQYQALLAYSGDKTDEGIRLMQSDYGNLSYHINKNKLFI  
FSASNDQAQNPNTLALFFYEDAKQGLITVAGDRSEKRFNGSNHCGITAMWCLISAP  
YEASVRFRTPIQIAGTSFAPITGTALLQKQYPMNSNDNLRTLLTAAQDIGAV  
GVDSKFWGLLDKAMKNGPAPFPFDPTADTKGTSDIAYSFRNDISGTTGLIKKGS  
LOQLGNNTYTGKTIIEGSLVLYGNKNSDMRVETKALIIYNGAASGSLNSDGIYVL  
ADTDSGANETVHLKGDLOLGGECTLYRLGLKLLKVDGTAMTGGKLYMSRCKAGYL  
NRTQRPVFLSNAKIGRDYSFFNIETDGLGLASLDSVEKTAGSEDTLSYYVRGNA  
ARTASAAHPAGLKHAEVGGSNLENLMVELDASESSATPETVETAAADRTDMPGI  
RPGYATFAAAAVGHANAADVRIFNLSAATVYADSTAAHADMQRRKLKAYSDGLDHN  
ATGLRVIAQTQDDGTWBOGGVEMKRGSTQTVGAATGENTTAATLGMHGSWTSE  
NSNAKTDISLEFAGIRHDAGDIGYLGFLFSYGRVKNISIRSTGADSHAEGSVNGTLM  
OLGALGVNVPFAATGDLTVBGGRLYDLKQDAFAEKGSALGWSGNSISITEGTLVGLAG  
LKLQPLSDKAVLFATAGVERDLNGRDYTVTGGTGATAATGKTGARNMPTRLVAGL  
GADVEFGNWGLARYISYAGSKQYGNHSGRVGVR"

BASE COUNT 830 a 909 c 900 g 565 t

ORIGIN

Query Match 100.0%; Score 3202.4; DB 1; Length 3204;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 3203; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	ATGCGAACACCCCAACCTTCCTCAAAAACCTTCAAAACCGGCTGCCATGGCGTTAGCT	60
DB	1	ATGCGAACACCCCAACCTTCCTCAAAAACCTTCAAAACCGGCTGCCATGGCGTTAGCT	60
QY	61	GTTCGCAACAACTTTCTGCCTGTCTTAGCGCGGGGGGGGCACTTCTCGCGCCGACTTC	120
DB	61	GTTCGCAACAACTTTCTGCCTGTCTTAGCGCGGGGGGGGCACTTCTCGCGCCGACTTC	120
QY	121	AATGACGGCGGACCGGTATCGGCAGCAACACAGACAGCAACACAGCAAAATCAGCAGCA	180
DB	121	AATGACGGCGGACCGGTATCGGCAGCAACACAGACAGCAACACAGCAAAATCAGCAGCA	180
QY	181	GTATCTTACCGCGGTATCAAGAACGAAATGTGCAAGACAGCAAGCATGCTGTGCCCGT	240
DB	181	GTATCTTACCGCGGTATCAAGAACGAAATGTGCAAGACAGCAAGCATGCTGTGCCCGT	240
QY	241	CGGGATGACGTTGCGGTTACAGACAGGATGCCAAAATCAATGCCCGCCCCCGGAATCTG	300
DB	241	CGGGATGACGTTGCGGTTACAGACAGGATGCCAAAATCAATGCCCGCCCCCGGAATCTG	300
QY	301	CATACCGGAGACTTTTCAAAACCCAAATGACGCATACAAGAAATTTGATCAACCTCAAACT	360
DB	301	CATACCGGAGACTTTTCAAAACCCAAATGACGCATACAAGAAATTTGATCAACCTCAAACT	360
QY	361	GCAATTGAAGCAGGCTATACGACGCGGGGTAGAGGTAGGTATCGTCGATACAGCGCAA	420
DB	361	GCAATTGAAGCAGGCTATACGACGCGGGGTAGAGGTAGGTATCGTCGATACAGCGCAA	420
QY	421	TCCGTCGGCAGCATATCCTTTCCGNACTGTATGCGCAGAAAAGAACACCGGCTATACGAA	480
DB	421	TCCGTCGGCAGCATATCCTTTCCGNACTGTATGCGCAGAAAAGAACACCGGCTATACGAA	480
QY	481	AATTACAAAACCTATACGCGGTATATGCGGAAGGAGCGCCTGAAGACGGAGCGGTAA	540
DB	481	AATTACAAAACCTATACGCGGTATATGCGGAAGGAGCGCCTGAAGACGGAGCGGTAA	540
QY	541	GACATTAAAGCTTCTTTTCGACGATGAGGCCGTTATAGAGACTGAAGCAAGCCGACGGAT	600
DB	541	GACATTAAAGCTTCTTTTCGACGATGAGGCCGTTATAGAGACTGAAGCAAGCCGACGGAT	600
QY	601	ATCCGCCACGTAAAGAAATCGGACACATCGATGGTCTCCCATATATTATTGGCGGCGT	660
DB	601	ATCCGCCACGTAAAGAAATCGGACACATCGATGGTCTCCCATATATTATTGGCGGCGT	660

DB	601	ATCCGCCACGTAAAGAAATCGGACACATCGATGGTCTCCCATATATTATTGGCGGCGGT	660
QY	661	TCCGTGGAGCGCAGACCTGCGAGCGGTATTGCGCCCGCATGCGACGCTACACATAATGAAT	720
DB	661	TCCGTGGAGCGCAGACCTGCGAGCGGTATTGCGCCCGCATGCGACGCTACACATAATGAAT	720
QY	721	ACGCATGATGAACAACGAAAGAAATAATGTCTGAGCCATCCGCAATGCGATGGGTCAAG	780
DB	721	ACGCATGATGAACAACGAAAGAAATAATGTCTGAGCCATCCGCAATGCGATGGGTCAAG	780
QY	781	CTGGCGGAACCTGGCGTCCGATCGTCAATACAGTGTGGAAACAACATCGAGGCGACGC	840
DB	781	CTGGCGGAACCTGGCGTCCGATCGTCAATACAGTGTGGAAACAACATCGAGGCGACGC	840
QY	841	ACTGCCGACCACTTTCCTCAATAGCCAAATTCGAGGAGCAGTACCGCAAGCGTTGCTCGCC	900
DB	841	ACTGCCGACCACTTTCCTCAATAGCCAAATTCGAGGAGCAGTACCGCAAGCGTTGCTCGCC	900
QY	901	TATTCGCGCGGTGATPAAACAGACGAGGATATCCGCCGTGATGCAACAGAGGATTTACGCG	960
DB	901	TATTCGCGCGGTGATPAAACAGACGAGGATATCCGCCGTGATGCAACAGAGGATTTACGCG	960
QY	961	RACTTGTCTCTACCACATCCGTAATAAAAACATGCTTTTCATTTTTTCGGCAGCAATGAC	1020
DB	961	RACTTGTCTCTACCACATCCGTAATAAAAACATGCTTTTCATTTTTTCGGCAGCAATGAC	1020
QY	1021	GCACAAGCTCAGCCCCACACACTGACCTATTGCCCATTATGAAAAAGATGCTCAAAA	1080
DB	1021	GCACAAGCTCAGCCCCACACACTGACCTATTGCCCATTATGAAAAAGATGCTCAAAA	1080
QY	1081	GGCATTATCACAGTCCGAGCGGTAGACCCGAGTGGAGAAAAAGTTCAATGGGTCCCAACCAT	1140
DB	1081	GGCATTATCACAGTCCGAGCGGTAGACCCGAGTGGAGAAAAAGTTCAATGGGTCCCAACCAT	1140
QY	1141	TCCGAAATTTACTGCCATCTGGTGCCTATCGGCACCTATGAGCAAGCGTCCGTTTCACC	1200
DB	1141	TCCGAAATTTACTGCCATCTGGTGCCTATCGGCACCTATGAGCAAGCGTCCGTTTCACC	1200
QY	1201	CGTACAAACCCCGATTCAAAATTCGCGGAACATCCTTTTCCGACCCCATGTAACCGGCAG	1260
DB	1201	CGTACAAACCCCGATTCAAAATTCGCGGAACATCCTTTTCCGACCCCATGTAACCGGCAG	1260
QY	1261	GGCGTCTGCTGCTCGAANAATACCCGTGGATGAGCAAGCAGCAACCTGGGTACACAGCTG	1320
DB	1261	GGCGTCTGCTGCTCGAANAATACCCGTGGATGAGCAAGCAGCAACCTGGGTACACAGCTG	1320
QY	1321	CTGACAACGGCTCAGACATCGTGCAGTCGCGGTGGACAGCAAGTTCGGGTGGGACTG	1380
DB	1321	CTGACAACGGCTCAGACATCGTGCAGTCGCGGTGGACAGCAAGTTCGGGTGGGACTG	1380
QY	1381	CTGGATGGGGTAAAGCCCATGAACGACCGCGTCTTTTCCGTTCCGGGACTTTACCGCC	1440
DB	1381	CTGGATGGGGTAAAGCCCATGAACGACCGCGTCTTTTCCGTTCCGGGACTTTACCGCC	1440
QY	1441	GATACGAAAGGTACATTCGATATTCGCTACTCTTCCGTAACGACATTTTCAGGCGAGGGC	1500
DB	1441	GATACGAAAGGTACATTCGATATTCGCTACTCTTCCGTAACGACATTTTCAGGCGAGGGC	1500
QY	1501	GGCCTGATCAAAAAGGCGGACCACTGCAACTGCAAGCAACAACACCTTATACGGGC	1560
DB	1501	GGCCTGATCAAAAAGGCGGACCACTGCAACTGCAAGCAACAACACCTTATACGGGC	1560
QY	1561	AAAAACCATTTATCGAAGGGGTTTCGCTGGTGTGTGTACGGCAACAACAAATTCGATATCGC	1620
DB	1561	AAAAACCATTTATCGAAGGGGTTTCGCTGGTGTGTGTACGGCAACAACAAATTCGATATCGC	1620
QY	1621	GTGCAAAACAAAGGTGCGCTGATTTTAAACGGGGGGCATCCGGCGGTAGCCTGAACAGC	1680
DB	1621	GTGCAAAACAAAGGTGCGCTGATTTTAAACGGGGGGCATCCGGCGGTAGCCTGAACAGC	1680
QY	1681	GACGCATTGTCTATCTGGCAGATACCGACCATCCGGCGCAACGAAACCGTGCACATC	1740
DB	1681	GACGCATTGTCTATCTGGCAGATACCGACCATCCGGCGCAACGAAACCGTGCACATC	1740

QY	2821	GACCTGCTCAAAACAGGATGCAATTCGCCGAAAAAGAGCAGTGCTTTGGCTTGAGCGGCAAC	2880
Db	2821	GACCTGCTCAAAACAGGATGCAATTCGCCGAAAAAGAGCAGTGCTTTGGCTTGAGCGGCAAC	2880
QY	2881	AGCCTCACTGAAGGCACACGTGGTCGGACATCGCGGGTCTGAAGCTCTGCGCAACCCCTTGAGC	2940
Db	2881	AGCCTCACTGAAGGCACACGTGGTCGGACATCGCGGGTCTGAAGCTCTGCGCAACCCCTTGAGC	2940
QY	2941	GATAAAGCCGTCTCTGTTTGTCAACGCGCGGGCGTGGAAACGCGACCTCAACGCGCGGACTAC	3000
Db	2941	GATAAAGCCGTCTCTGTTTGTCAACGCGCGGGCGTGGAAACGCGACCTCAACGCGCGGACTAC	3000
QY	3001	ACGGTAACGGGCGGCTTTTACCGCGCGGACTGACGACAACCGGCAAGACGGGGCACGCAAT	3060
Db	3001	ACGGTAACGGGCGGCTTTTACCGCGCGGACTGACGACAACCGGCAAGACGGGGCACGCAAT	3060
QY	3061	ATGCCGCACACCCGCTGTTGCCGCTCTGCGCGGGATGTCGAATTCGGCAACGGCTGG	3120
Db	3061	ATGCCGCACACCCGCTGTTGCCGCTCTGCGCGGGATGTCGAATTCGGCAACGGCTGG	3120
QY	3121	AACGGCTTGGCAGCTTACAGCTACGCGGTTCCAAACAGTACGCGCAACACACGCGGACGA	3180
Db	3121	AACGGCTTGGCAGCTTACAGCTACGCGGTTCCAAACAGTACGCGCAACACACGCGGACGA	3180
QY	3181	GTCGGCGTAGGCTACCGGTTCTGA	3204
Db	3181	GTCGGCGTAGGCTACCGGTTCTGA	3204
RESULT 3			
LOCUS	NMA222491/c		
DEFINITION	Neisseria meningitidis serogroup A strain Z2491 complete genome	DNA	linear BCT 04-DEC-
ACCESSION	segment 2/7		
VERSION	AL162753	AL157959	
KEYWORDS	AL162753.2	GI:7379120	
SOURCE	Neisseria meningitidis Z2491.		
ORGANISM	Neisseria meningitidis Z2491.		
REFERENCE	1 (bases 1 to 349061)		
AUTHORS	Parkhill,J., Achtman,M., James,K.D., Bentley,S.D., Churcher,C., Klee,S.R., Morelli,G., Basham,D., Brown,D., Chillingworth,T., Davies,R.M., Davis,P., Devlin,K., Feltwell,T., Hamlin,N., Holroyd,S., Jagels,K., Leather,S., Moule,S., Mungall,K., Quail,M.A., Rajandream,M.A., Rutherford,K.M., Simmonds,M., Skelton,J., Whitehead,S., Spratt,B.G. and Barrell,B.G.		
TITLE	Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491		
JOURNAL	Nature 404 (6777), 502-506 (2000)		
MEDLINE	20222556		
PUBMED	10761919		
REFERENCE	2 (bases 1 to 349061)		
AUTHORS	Parkhill,J.		
TITLE	Direct Submission		
JOURNAL	Submitted (30-MAR-2000) Submitted on behalf of the Neisseria sequencing team, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk		
COMMENT	Details of N. meningitidis sequencing at the Sanger Centre are available on the World Wide Web. (URL, <a href="http://www.sanger.ac.uk/Projects/N_meningitidis/">http://www.sanger.ac.uk/Projects/N_meningitidis/</a> ).		
FEATURES	Location/Qualifiers		
source	1. .349061 /organism="Neisseria meningitidis Z2491" /strain="Z2491" /db_xref="taxon:122587" /note="serogroup: A" complement(24. .206) /note="ATR repeat; hmms hit to HMM ATR (1 - 183), score 310.39" /label=ATR		
repeat_unit			

```

RBS      209..212
gene      /gene="NMA0368"
CDS       220..681
          /gene="NMA0368"
          /note="NMA0368, probable integral membrane protein, len:
153 aa; contains four probable transmembrane domains"
          /codon_start=1
          /transl_table=11
          /product="putative integral membrane protein"
          /protein_id="CAB83669.1"
          /db_xref="GI:7379121"
          /db_xref="SPTREMBL:Q9JWH7"
          /translation="MQEQRKSSFPVIMLLVSLVALWIASLNVAFYLGNGHSMGLTV
LILGISFASLDIORYCAVYANVYVWIAAIVLLALRKKVPDVHAAFWGLALVAFSVKAVYV
DEAGNTSDIVRYGAGFYLYWAAFAVASIGTITAGNKKERKKAASADGTNDV"
          671..1492
          /gene="hemk"
          671..1492
          /gene="hemk"
          /note="NMA0369, hemk, Hemk protein, len: 273 aa; similar
to e.g. HEMK_ECOLI P37186 HEMK protein (277 aa); fasta
scores; E(): 0, 42.3% identity in 279 aa overlap. Contains
PS00092 N-6 Adenine-specific DNA methylases signature"
          /codon_start=1
          /transl_table=11
          /product="Hemk protein"
          /protein_id="CAB83670.1"
          /db_xref="GI:7379122"
          /db_xref="SPTREMBL:Q9JWH6"
          /translation="MTFDKWLGLSKLPKNEARMLQYVSEYTRVOLLTRGGEPMPEV
RQRADLRQRLNPEVAYILGAREFYRFTVNPVSILIPRTEHLVAVLARLPEN
GRVWDLGSGAVATVALERPDFAFVRASDLSPPALETARKNAADLGARVFAVYGSWF
DTDMPSGKWDIIIVSNPPYIENGDKHLSQGLRPEPOIALTFDSGLSCITLAQGAP
DRLAEGGFLLEHGFDQGAARVGLAENGFSGVETLPDLGLDRVTLGKYNKHLK"
          831..840
          /gene="hemk"
          /note="Core DNA uptake sequence: gccgtctgaa"
          /label=DUS
          1162..1171
          /gene="hemk"
          /note="Core DNA uptake sequence: gccgtctgaa"
          /label=DUS
          1187..1207
          /gene="hemk"
          /note="PS00092 N-6 Adenine-specific DNA methylases
signature"
          complement(1279..1288)
          /note="Core DNA uptake sequence: gccgtctgaa"
          /label=DUS
          1565..1568
          1578..2966
          /gene="NMA0370"
          1578..2966
          /gene="NMA0370"
          /note="NMA0370, probable integral membrane protein, len:
462 aa; similar to hypothetical proteins e.g. Y325_HAEIN
P44640 hypothetical protein HI0325 (450 aa), fasta scores;
E(): 0, 49.8% identity in 464 aa overlap"
          /codon_start=1
          /transl_table=11
          /product="putative integral membrane protein"
          /protein_id="CAB83671.1"
          /db_xref="GI:7379123"
          /db_xref="SPTREMBL:Q9JWH5"
          /translation="MNAVVAIVIMVLVLSLRVHVLLTTCFAFVGGAAGMPLONTA
DAAGQVQAGIIPVFNKLEGGAIALSYAMIGAFAMAITHSGLPQQLAGAVVRKLRN
GCMGPSVRSRGVVKWLLLSILVNGMMSQNIPTHTAFIPMIVPPLLVLVFNKLRIDR
RLIACVITFLVTTYMFLPYGFGAIFLNEILLGNHSAAPOLDVKNNVMAAIPAL
GMLAGLLAFVHYRKPRLYQSNADTAGNDAANRPOPASVRSIAAAVAJAVCFAL
MYESLVLGALMGFAVFMVLGVINRDKANDVFEGIKMMAMVGFIMIAAQGAAMNA
TGHQIPLVSSMAIFGNSKGMMAALMLVGLLVMTGIGSSFSTLPTITAAIYVPLCVGL
GFSPLATVAIVGTAGALGDAGSPASDSTLGPTMGLNADGQHDHIRDSTIPTFIHYNIP

```

```

LLIAGWIAAMVL"
2967..3191
/gene="slyx"
2967..3191
/gene="slyx"
/note="NMA0371, slyx, SLX protein homolog, len: 74 aa;
similar to SLX_HAEIN P44759 SLX protein homolog (73 aa),
fasta scores; E(): 0.0017, 33.8% identity in 74 aa
overlap, and SLX_ECOLI P30857 SLX protein (72 aa), fasta
scores; E(): 0.24, 32.4% identity in 68 aa overlap"
/codon_start=1
/transl_table=11
/product="SLX protein homolog"
/protein_id="CAB83672.1"
/db_xref="GI:7379124"
/db_xref="SPTREMBL:Q9JWH4"
/translation="MDAVQEFHEHRITELEIGISALQEDVIAGLNAMVAELRQTLQQA
QLRLLYQRMQDRNPDAOEYPYSLRDEIIPHY"
complement(3271..3423)
/gene="NMA0372"
complement(3271..3423)
/gene="NMA0372"
/note="NMA0372, unknown, questionable CDS, len: 50 aa"
/codon_start=1
/transl_table=11
/product="very hypothetical protein NMA0372"
/protein_id="CAB83673.1"
/db_xref="GI:7379125"
/db_xref="SPTREMBL:Q9JWH3"
/translation="MFGRLSPGSDGFFLAVYEPVGFICLOGKTAYPYRSGGKRPFHRI
KGIUSD"
3562..3571
/note="Core DNA uptake sequence: gccgtctgaa"
/label=DUS
complement(3586..4356)
/gene="thif"
complement(3586..4356)
/gene="thif"
/note="NMA0373, thif, probable Thif protein, len: 256 aa;
similar to e.g. THIF_ECOLI P30138 THIF protein (251 aa),
fasta scores; E(): 0, 43.1% identity in 246 aa overlap,
and MOEB_ECOLI P12282 molybdopterin biosynthesis MOEB
protein. (249 aa), fasta scores; E(): 0, 43.9% identity in
244 aa overlap (note that N.m. does not have orthologs of
any other molybdopterin biosynthesis proteins). Contains
pfam match to entry PF00899 Thif_family, Thif family"
/codon_start=1
/transl_table=11
/product="Thif protein"
/protein_id="CAB83674.1"
/db_xref="GI:7379126"
/db_xref="SPTREMBL:Q9JWH2"
/translation="MTTEHDNDADFLLRYSRHILLDEIGIEGQOKLSAAHILVVGCG
GIGAAALPYLAASGITITDIADSTVELHNLQRFVAFDEGVDVKLTALADRLRHIN
HTDVVRTINEKLDGCRITGLVQAAADIVLDCDNDATRVQAVNRACVQAKTPLYSGAAR
FEGVLAVYRDLDPSPCYACLFDDGSGADSLGFSPLVGFSPVLTIGTGQAALAILL
DAGPSHGLRAYRALEGGWQYFDLPRNPECPVCGAER"
3787..3796
/note="Core DNA uptake sequence: gccgtctgaa"
/label=DUS
complement(3847..4257)
/gene="thif"
/note="Pfam match to entry PF00899 Thif_family, Thif
family, score 186.60, E-value 4.1e-52"
complement(4418..4427)
/note="Core DNA uptake sequence: gccgtctgaa"
/label=DUS
4471..7224
/gene="ppc"
4471..7224
/gene="ppc"
/EC_number="4.1.1.31"
/note="NMA0374, ppc, phosphoenolpyruvate carboxylase, len:

```

917 aa; similar to many e.g. CAPP\_RHOPA O32483  
phosphoenolpyruvate carboxylase (EC 4.1.1.31) (936 aa),  
fasta scores: E(): 0, 43.3% identity in 928 aa overlap.  
Contains 2x Pfam match to entry PF00311 PEPCase.  
Phosphoenolpyruvate carboxylase, PS00017 ATP/GTP-binding  
site motif A (P-loop), and PS00393 Phosphoenolpyruvate  
carboxylase active site 2  
/codon\_start=1  
/transl\_table=11  
/product="phosphoenolpyruvate carboxylase"  
/protein\_id="CA883675.1"  
/db\_xref="GI:7379127"

Query Match		98.6%;	Score 3160.6;	DB 1;	Length 349061;	
Best Local Similarity		99.3%;	Pred. No. 0;			
Matches 3185;		Conservative	0;	Mismatches	19;	Indels 3; Gaps 1;
QY	1	ATCGGAACGACCCACCTTCCCTACAAAACCTTTCAAACCGGCTGCCATGGCGTTAGCT	60			
Db	126320	ATCGGAACGACCCCAACCTTCCCTACAAAACCTTTCAAACCGGCTGCCATGGCGTTAGCT	126261			
QY	61	GTTCGCAACACACTTCTGCTGCTTAGGCGCGCGC---GGCGGCATTTCTGGCCCGCAG	117			
Db	126260	GTTCGCAACACACTTCTGCTGCTTAGGCGCGCGGAGGCGGCACTTCTGGCCCGCAG	126201			
QY	118	TTCAAATGCGAGCGGCACCGGTATCGGCAGCAACAGACAGCAACACAGCGAAATCAGCA	177			
Db	126200	TTCAAATGCGAGCGGCACCGGTATCGGCAGCAACAGACAGCGGCAACAGCGAAATCAGCA	126141			
QY	178	GCAGTATCTTACGCCCGGTATCAAGAACGAAATGTGCAAGACAGAGCATGCTCTGTGCC	237			
Db	126140	GCAATATCTTACGCCCGGTATCAAGAACGAAATGTGCAAGACAGAGCATGCTCTGTGCC	126081			
QY	238	GGTCGGGATGACGTTGGGTTACAGACAGGGATGCCAAATCAATGCCCCCCCCCCGAAT	297			
Db	126080	GGTCGGGATGACGTTGGGTTACAGACAGGGATGCCAAATCAATGCCCCCCCCCCGAAT	126021			
QY	298	CTGCATACCGGAGACTTTACAAACCCCAATGAGGCATACAGAATTTGATCAACCTCAAA	357			
Db	126020	CTGCATACCGGAGACTTTACAAACCCCAATGAGGCATACAGAATTTGATCAACCTCAAA	125961			
QY	358	CTTGCAATTTAAGCAGGCTATACAGACGCGGGTAGAGTAGGTATGTCGTATACAGGC	417			
Db	125960	CTTGCAATTTAAGCAGGCTATACAGACGCGGGTAGAGTAGGTATGTCGTATACAGGC	125901			
QY	418	GAATCGTCGGCAGCATATCTTTCCGGAACGTGATGGCAAGAAAGAACACGCGTATAAC	477			
Db	125900	GAATCGTCGGCAGCATATCTTTCCGGAACGTGATGGCAAGAAAGAACACGCGTATAAC	125841			
QY	478	GAAATTTACAAAACATATACGGCGGTATATCGGGAAGGAAGCGCTGAAGACGGAGCGGT	537			
Db	125840	GAAATTTACAAAACATATACGGCGGTATATCGGGAAGGAAGCGCTGAAGACGGAGCGGT	125781			
QY	538	AAGACATTTAAGCTTTCTTTGACAGTAGGCGGTTATAGAGACTGAAGCAAAAGCCGACG	597			
Db	125780	AAGACATTTAAGCTTTCTTTGACAGTAGGCGGTTATAGAGACTGAAGCAAAAGCCGACG	125721			
QY	598	GATATCGGCCACCTAAAGAAATCGGACACATCGATGTGCTCTCCCATATATTGGCGGG	657			
Db	125720	GATATCGGCCACCTAAAGAAATCGGACACATCGATGTGCTCTCCCATATATTGGCGGG	125661			
QY	658	CGTTCGCTGGACGGCAGACCTCGAGCGGTTATTTGGCGCGGATCGGACGCTACACATAATG	717			
Db	125660	CGTTCGCTGGACGGCAGACCTCGAGCGGTTATTTGGCGCGGATCGGACGCTACACATAATG	125601			
QY	718	AATACGCATGATGGAACCAAGAACGAAATATGTCTGACGCCATCCGCAATGATGGGTC	777			
Db	125600	AATACGCATGATGGAACCAAGAACGAAATATGTCTGACGCCATCCGCAATGATGGGTC	125541			
QY	778	AACTGGGCAACGTCGCGTCGATCGTCAATAACAGTTTTTGGAAACAAATCGAGGGCA	837			
Db	125540	AACTGGGCAACGTCGCGTCGATCGTCAATAACAGTTTTTGGAAACAAATCGAGGGCA	125481			

QY	838	GGCACTGCCGACCACTTTCCAAATAGCCAATTCGGAGGAGCAGTACCGCCCAAGCGTTGCTC	897			
Db	125480	GGCACTGCCGACCACTTTCCAAATAGCCAATTCGGAGGAGCAGTACCGCCCAAGCGTTGCTC	125421			
QY	898	GCCTATTTCCGGCGGTGATAAACAAGACGAGGGTATCCGCTGATGCAACAGAGCGATTAC	957			
Db	125420	GCCTATTTCCGGCGGTGATAAACAAGACGAGGGTATCCGCTGATGCAACAGAGCGATTAC	125361			
QY	958	GGCAACTTGTCTTACCACATCCGTAATAAACAATGCTTTTCATTTTTCGGCAAGCAAT	1017			
Db	125360	GGCAACTTGTCTTACCACATCCGTAATAAACAATGCTTTTCATTTTTCGGCAAGCAAT	125301			
QY	1018	GAGCACAAGCTCAGCCCAACACACTGACCTATTGGCATTTTATGAAAAGATGCTCAA	1077			
Db	125300	GAGCACAAGCTCAGCCCAACACACTGACCTATTGGCATTTTATGAAAAGATGCTCAA	125241			
QY	1078	AAGGCAATTATCACAGTCGCGAGGCTAGACCGCAGTGGAGAAAAGTTCAATGGCTCCAAC	1137			
Db	125240	AAGGCAATTATCACAGTCGCGAGGCTAGACCGCAGTGGAGAAAAGTTCAATGGCTCCAAC	125181			
QY	1138	CATTGCGGAATTAAGTGCATGCTGGTGCCTATCCGACCCCTATGAAGCAAGCGTCCGTTTC	1197			
Db	125180	CATTGCGGAATTAAGTGCATGCTGGTGCCTATCCGACCCCTATGAAGCAAGCGTCCGTTTC	125121			
QY	1198	ACCGTACAAACCCGATTCAAAATTCGCGGAACATCTTTCCGACCCATCTGTAACCGGC	1257			
Db	125120	ACCGTACAAACCCGATTCAAAATTCGCGGAACATCTTTCCGACCCATCTGTAACCGGC	125061			
QY	1258	ACGCGGCTCTGCTGCTGTCAGAAATACCCGTGATGAGCAACACACCTGCGTACCACG	1317			
Db	125060	ACGCGGCTCTGCTGCTGTCAGAAATACCCGTGATGAGCAACACACCTGCGTACCACG	125001			
QY	1318	CTGCTGACACGGCTCAGGACATCGGTGCGCTGCGCTGACACCAAGTTTCGGCTGGGA	1377			
Db	125000	CTGCTGACACGGCTCAGGACATCGGTGCGCTGCGCTGACACCAAGTTTCGGCTGGGA	124941			
QY	1378	CTGCTGATCGGGTAAGGCGCATGAACGACCCGCTCTTTCCGTCGGCGACTTTACC	1437			
Db	124940	CTGCTGATCGGGTAAGGCGCATGAACGACCCGCTCTTTCCGTCGGCGACTTTACC	124881			
QY	1438	CGCGATACGAAAGTACATCGGATATTGCTTCTCTCGTACGACATTTTCAGGCACG	1497			
Db	124880	CGCGATACGAAAGTACATCGGATATTGCTTCTCTCGTACGACATTTTCAGGCACG	124821			
QY	1498	GGCGGCTGATCAAAAAGCGCGAGCAACTGCACTGACGCGCAACACACCTATACG	1557			
Db	124820	GGCGGCTGATCAAAAAGCGCGAGCAACTGCACTGACGCGCAACACACCTATACG	124761			
QY	1558	GGCAAAACCACTTATCGAAGCGGTTCTGCTGGTTGTGTACGGCAACAACATCGGATATG	1617			
Db	124760	GGCAAAACCACTTATCGAAGCGGTTCTGCTGGTTGTGTACGGCAACAACATCGGATATG	124701			
QY	1618	CGGCTCGAAACCAAGTGGCTGATTTATAACGGGGGGGATCCGGGGTAGCCTGAAC	1677			
Db	124700	CAGTCTGAAACCAAGTGGCTGATTTATAACGGGGGGGATCCGGGGTAGCCTGAAC	124641			
QY	1678	AGCGACGGCTATCTCTATCTGGCAGATACCGACCGATCCGCGCAACAGCAACCGTGCAC	1737			
Db	124640	AGCGACGGCTATCTCTATCTGGCAGATACCGACCGATCCGCGCAACAGCAACCGTGCAC	124581			
QY	1738	ATCAAGGGCGATCTGACGCTGGCGCGGAAGTACGCTGTACACACGTTTGGGCAAACTG	1797			
Db	124580	ATCAAGGGCGATCTGACGCTGGCGCGGAAGTACGCTGTACACACGTTTGGGCAAACTG	124521			
QY	1798	CTGAAAGTGGACGGTACGGCGATGACCGGGGAAGGTGTACATGTGCGACGCGGCAAA	1857			
Db	124520	CTGAAAGTGGACGGTACGGCGATGACCGGGGAAGGTGTACATGTGCGACGCGGTAAG	124461			
QY	1858	GGGCGGCTATCTCAACCGTACCGGACACGCTGCTCCCTCCCTGAGTCCGCCAATC	1917			
Db	124460	GGGCGGCTATCTCAACCGTACCGGACACGCTGCTCCCTCCCTGAGTCCGCCAATC	124401			
QY	1918	GGGCGGATATTCTTTCTTACAAAACATCGAAACCGAGCGGTGCTGCTGCTTCCCTC	1977			

|||||  
Db 124400 GGGCGGGATTATTCCTTCTTCAACAACATCGAAACCGACGGCGGTCTGTCGTTCCCTTC 124341  
QY 1978 GACAGCGTCGAAAAACACGCGGCGAGTGAAGCGGCACACGCTGCTCCTATTATTCCTCGTCCG 2037  
|||||  
Db 124340 GACAGCGTCGAAAAACACGCGGAGTGAAGCGGCACACGCTGCTCCTATTATTCCTCGTCCG 124281  
QY 2038 GGCAATGGCGGCACGAGCTGCTTCGCGACGGGCACATTCGCGCGCGCGCGGTCTGAAACAC 2097  
Db 124280 GGCAATGGCGGCACGAGCTGCTTCGCGACGGGCACATTCGCGCGCGCGCGGTCTGAAACAC 124221  
QY 2098 GCCGTAGAACAGGGGGCGAGCAATCTGGAACACCTGATGGTGCAGACTGATGCTCCGAA 2157  
|||||  
Db 124220 GCCGTAGAACAGGGGGCGAGCAATCTGGAACACCTGATGGTGCAGACTGATGCTCCGAA 124161  
QY 2158 TCATCCGCAACACCGAGACGCTTGAACCTGCGGCGCGCGACGCGCACATATGCGGGCG 2217  
Db 124160 TCATCCGCAACACCGAGACGCTTGAACCTGCGGCGCGCGACGCGCACATATGCGGGCG 124101  
QY 2218 ATCCGCCCCCTACGGCGCAACTTTCGCGCGAGCGGCGATACAGCATGCGAATGCGCGC 2277  
Db 124100 ATCCGCCCCCTACGGCGCAACTTTCGCGCGAGCGGCGATACAGCATGCGAATGCGCGC 124041  
QY 2278 GACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCGCC 2337  
|||||  
Db 124040 GACGGTGTACGCATCTTCAACAACTTCGCGCGCTACCGTCTATGCCGACAGTACCGCGCC 123981  
QY 2338 CATGCCGATATGACGGAGCGCGGCTGAAAGCGGTATCGAGCGGTTGACCAACACGCT 2397  
Db 123980 CATGCCGATATGACGGAGCGCGGCTGAAAGCGGTATCGAGCGGTTGACCAACACGCT 123921  
QY 2398 ACGGGTCTCGCGCTCATCGCGCAACCCCAACAGGAGCGGTGGAACGTTGGAACAGGGCGGT 2457  
Db 123920 ACGGGTCTCGCGCTCATCGCGCAACCCCAACAGGAGCGGTGGAACGTTGGAACAGGGCGGT 123861  
QY 2458 GTTGAAGGCAAAATGCGCGGCGAGTACCCAAACCGTCTGCGCATTCGCCGCAAAACCGCGGAA 2517  
|||||  
Db 123860 GTTGAAGGCAAAATGCGCGGCGAGTACCCAAACCGTCTGCGCATTCGCCGCAAAACCGCGGAA 123801  
QY 2518 AATACACAGCAGCGCCACACTGGGCATGAGGACACAGCAGCATGAGCGCAAAACAGTGCA 2577  
Db 123800 AATACACAGCAGCGCCACACTGGGCATGAGGACACAGCAGCATGAGCGCAAAACAGTGCA 123741  
QY 2578 AATGCAAAACCGACAGCATATGCTGTTTGAGGCGATACGGACAGATCGGGCGGATATC 2637  
Db 123740 AATGCAAAACCGACAGCATATGCTGTTTGAGGCGATACGGACAGATCGGGCGGATATC 123681  
QY 2638 GCGTATCTCAAGGCGCTGTTCTCTACGACGCTACAAAACAGCATCAGCGCAGCACCC 2697  
Db 123680 GCGTATCTCAAGGCGCTGTTCTCTACGACGCTACAAAACAGCATCAGCGCAGCACCC 123621  
QY 2698 GGTGCGGAGCAACATCGGGAAGCAGCGTCAACGGCACGCTGATGACGCTGGGCGGCACTG 2757  
Db 123620 GGTGCGGAGCAACATCGGGAAGCAGCGTCAACGGCACGCTGATGACGCTGGGCGGCACTG 123561  
QY 2758 GGGGTGTCAAGCTTCGTTTCCGCAACGGGAGATTTGACGGTCAAGGCGGTCTGCGC 2817  
Db 123560 GGGGTGTCAAGCTTCGTTTCCGCAACGGGAGATTTGACGGTCAAGGCGGTCTGCGC 123501  
QY 2818 TAGCACTCTCAACAGGATGATTCGCGCAAAAGGAGTCTTTGGGCTGGAGCGGC 2877  
Db 123500 TAGCACTCTCAACAGGATGATTCGCGCAAAAGGAGTCTTTGGGCTGGAGCGGC 123441  
QY 2878 AACGAGCTCTCAAGGCACTGTTGCGGACTCGCGGGTCTGAAGCTGTCGCAACCCCTTG 2937  
Db 123440 AACGAGCTCACTGAAGGCACTGTTGCGGACTCGCGGGTCTGAAGCTGTCGCAACCCCTTG 123381  
QY 2938 AGCGATAAGCCGCTCTGTTTCAACAGGGCGGCGTGGAGCGGACCTGAACGAGCGCGAC 2997  
Db 123380 AGCGATAAGCCGCTCTGTTTCAACAGGGCGGCGTGGAGCGGACCTGAACGAGCGCGAC 123321  
QY 2998 TACACGGTACGGCGGCTTTACCGCGCGGCGACTGACGACACCGGCAAGCGGGGCGACGC 3057  
|||||

Db 123320 TACACGGTAACGGGCGGCTTTTACCGCGCGGACTCGAGCAACCGGCAAGACGGGGCGCACGC 123261  
QY 3058 AATATSCCGGCACACCCGCTGTTGTCGCGGTCTGGCGGGGATGTGAAATTCGGCAACGC 3117  
|||||  
Db 123260 AATATSCCGGCACACCCGCTGTTGTCGCGGTCTGGCGGGGATGTGAAATTCGGCAACGC 123201  
QY 3118 TGGACAGGCTTGGCACGTTTACAGCTACGCTACGCGGTTTCAAAACAGTACGGCAACACAGCGGA 3177  
Db 123200 TGGACAGGCTTGGCACGTTTACAGCTACGCTACGCGGTTTCAAAACAGTACGGCAACACAGCGGA 123141  
QY 3178 CGAGTCGCGGTAGGCTACCGGTTCTGA 3204  
Db 123140 CGAGTCGCGGTAGGCTACCGGTTCTGA 123114  
  
RESULT 4  
AE002545  
LOCUS Neisseria meningitidis serogroup B strain MC58 linear BCT 25-MAY-2000  
DEFINITION of the complete genome.  
ACCESSION AE002545 AE002098  
VERSION AE002545.1 GI:7227229  
KEYWORDS  
SOURCE  
ORGANISM Neisseria meningitidis MC58.  
Neisseria meningitidis MC58  
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;  
Neisseria.  
REFERENCE  
1 (bases 1 to 10591)  
Tettelin, H., Saunders, N.J., Heidelberg, J., Jeffries, A.C.,  
Nelson, K.E., Eisen, J.A., Ketchum, K.A., Hood, D.W., Peden, J.F.,  
Doudson, R.J., Nelson, W.C., Gwinn, M.L., DeBoy, R., Peterson, J.D.,  
Hickey, E.K., Haft, D.H., Salzberg, S.L., White, O., Fleischmann, R.D.,  
Dougherty, B.A., Mason, T., Ciecko, A., Parksey, D.S., Blair, E.,  
Cittone, H., Clark, E.B., Cotton, M.D., Utterback, T.R., Khouri, H.,  
Qin, H., Vamathevan, J., Gill, J., Scarlato, V., Masignani, V.,  
Piazza, M., Grandi, G., Sun, L., Smith, H.O., Fraser, C.M., Moxon, E.R.,  
Rappuoli, R. and Venter, J.C.  
Complete genome sequence of Neisseria meningitidis serogroup B  
strain MC58  
Science 287 (5459), 1809-1815 (2000)  
10717555  
20175755  
10710307  
REFERENCE  
2 (bases 1 to 10591)  
Tettelin, H., Saunders, N.J., Heidelberg, J., Jeffries, A.C.,  
Nelson, K.E., Eisen, J.A., Ketchum, K.A., Hood, D.W., Peden, J.F.,  
Doudson, R.J., Nelson, W.C., Gwinn, M.L., DeBoy, R., Peterson, J.D.,  
Hickey, E.K., Haft, D.H., Salzberg, S.L., White, O., Fleischmann, R.D.,  
Dougherty, B.A., Mason, T., Ciecko, A., Parksey, D.S., Blair, E.,  
Cittone, H., Clark, E.B., Cotton, M.D., Utterback, T.R., Khouri, H.,  
Qin, H., Vamathevan, J., Gill, J., Scarlato, V., Masignani, V.,  
Piazza, M., Grandi, G., Sun, L., Smith, H.O., Fraser, C.M., Moxon, E.R.,  
Rappuoli, R. and Venter, J.C.  
Direct Submission  
Submitted (17-MAR-2000) The Institute for Genomic Research, 9712  
Medical Center Dr, Rockville, MD 20850, USA  
  
FEATURES  
source  
1. 10591  
/organism="Neisseria meningitidis MC58"  
/db\_xref="taxon:122586"  
/strain="MC58"  
/note="serogroup: B"  
548. .3796  
/gene="NMB1969"  
548. .3796  
/gene="NMB1969"  
/note="similar to SP:P31631 GB:U07788 GB:M62363 PID:507431  
percent identity: 48.64; identified by sequence  
similarity: putative"  
/codon\_start=1  
/transl\_table=1  
/product="serotype-1-specific antigen, putative"  
/protein\_id="AAF42298.1"  
/db\_xref="GI:7227230"

/translation="MRTTPTPTKTKPTAMALAVATLTSACLCGGGGGTGAPDFNAG  
GTGSGSRATYAKSAAYSYAGIKNEMCKDRSMLCAGRDVAVTDRDAKINAPPNHLH  
TGDFPNDAVYKLNILNKPALPAEAGYTGVEVIGVDTGESVGSISFPELGRKEHYN  
ENYKNTAYMRKEAPEDGGDKDIEASFDEDAVIEAETPTDIRHVKETGIDLVSII  
GGRSDGPAGGIAPADPLHIMNTNDETKNEMVAARIANWVKLGERGVRIVNSFGT  
TSRAGTADLFQIANSEEQYRQALLDYSGDKTDEIRLMQWSDYLSHURKNLFI  
IFSTGNDAQNPATYALLPFYFEKDAQKGIITVAGVDRSGERFKRMEYGEPTPELEY  
SNHGCIITAMWLSIPYASVFRFTNPQIAGTSFSAPIVTGTAALLQKYPWMSNDN  
LRTTLLTQAQIDGAVGDSKFCWGLLDAGKAMNGPASFPGFTADTKGTSDIAYSFR  
NDISGTGLIKKGSQLOLHGNNTVTGKTIIEGSLVLYGNKNSDMRVEYTGALING  
AASGSLNSDGIYVLADTQDSANETVHIKSLQLDGKGLYTRLCGLKLLKVDGTALIG  
KGLTMSRKGAGNYNTRGRVPFLUSAKIGODYSFFNIETDGLGLSLASVETKAG  
SEGDLYSYVRGNNAARTASAAHAPAGLKHAVEQSGNSLENLMEVLDASSETAPE  
TVETAADRTDMPGIRPYQAFRAAAVQHANAADVRFNSLAATVYADSTAHAADM  
QGRLLKAGDLHDNGTGLRVIAOTQODGTWEOGVGEGKMRGTSOTVIAKTGCT  
TAAATLGMGRSTWBSNANAKTDSISLFLAGIRHDAGDGLYKGLFSYGRYKNSISRST  
GADEHAEGSVNGTLMQALGALGVNVPFAATGDLTVEGGLRVDLLKQDAFAEKGSLGW  
SGNSUTEGTLVGLAGLKUSQPLSKAVIFATAGVERDINGRDYITVVGFTGATAATGK  
TGARNMPTRLVAGLIGADVEFGNGWNGLIARYSYAGSKOYGNHSGRVGVGRF"

gene  
CDS

/db\_xref="GI:7227231"  
/gene="NMB1970"  
/note="similar to GB:AE000511 PID:2313385 percent identity: 53.22; identified by sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="para-aminobenzoate synthetase component 1/4-amino-4-deoxychorismate lyase, putative"  
/protein\_id="AAF42299.1"  
/db\_xref="GI:7227231"  
/translation="MPYFALPDDAVSGRAKRYONHVESRFRPELDALDGLQSGWO  
KGLHSLPADYGFGLPLGVSESGNLALEHFWANCADIDAESHLARSDGLPAGIST  
PQPSVSETDIDRIQIHEALRGTQIYNTIRHLQAYGNPVSIRLRKQVPIAY  
LSHPDAQSGSAWTLFCSPFLFKTGSDGTISTEPMKGTAPILGDQDERRAEALQAD  
PKNRAENYVLDLNDLKGIAQTGCVPEPKVSRFGSVQMWTSTIOAALPHTSF  
ADILRAEPGCSITGAPKMSMQIIESLEAEARGLYTGSIGLYNPSGGLGPEGFET  
VIRTLSPDSGIVGYGVGSGVIDSDPAEYREGCKARFLNELRDPGIFETL  
RAENGRCTLRHLRCLKTSQAALNLPDCENOIKOYLADLPGARFRVALLASDG  
ISLSRAVLNLRDKORVILISPAVLPAQNYLRFRFTICRALFDQANQWTAETQAFNSLF  
FNSDGLILEGGRSVYIRKNGWLTPSLDLDLILNGIMRQAVLDEPKYLIQTNQVLETH  
ITQKTLQAEIRLSNALRGVFAALA"  
complement(6668. .8185)  
/gene="NMB1971"  
complement(6668. .8185)  
/gene="NMB1971"  
/note="conserved hypothetical protein; identified by Glimmer2; putative"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="AAF42300.1"  
/db\_xref="GI:7227232"  
/translation="MLYFRYGLVVMCAAGVSAAYGADAPAILDKALLQVRSYSDK  
WAESDWKVENDPAPRVVDGDFLIAHPKMLEHSLRDALNGNQADLIAASLADLAKLPDYD  
AVLYGRALLAKLAPAEAVARYRELHGENAADERILLDLAAEFDDPFLKSAERH  
FAEAKLDPAPLVENGRFPRKKTGELTGWRFSGGISPAVNRNANNAAPQYCRONGRA  
QICVSRAERAGLNEYAEKLTPLADNHYLLFRSNIGGTSYIFFSKKSAYDDGFGRA  
YLWQYKNARQATGLPFYVOLSGSDGFDKTRKVRNNRLLPPYMLAHGVQVLSHTY  
RPNQWQFSALEHYRQRYQDRAEYNNRQDGFYVSAKRLGESATVFGWQVRF  
VPKRTGVGVNNAAYRRNYVAGWQEWROLGNSRVSYASARNTKGIAAFSTE  
QRNRMWVNSLASHDKLSYKGVIPALNYRFGRTESNVYAKRRNSEVVFVSADWRF"  
complement(8881. .10515)  
/gene="NMB1972"  
complement(8881. .10515)  
/gene="NMB1972"  
/note="similar to SP:P29842 GB:223008 GB:223009 PID:438185 PID:438187 percent identity: 99.26; identified by sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="chaperonin, 60 kDa"

/protein\_id="AAF42301.1"  
/db\_xref="GI:7227233"  
/translation="NAAKDVOFGNEVRKQWNVNLIANAVRTILGPKGRNVVVDRAF  
GGPHITKQGVTVAKETELKDFENMGAOMVKEASKTNDVAGDGTTTATVLAOSIVAE  
GNKYVTAGNPTDLARGIDKAVAALDELKNQKPCDTSKEIATQVGSISANSDQVGA  
IIAEAMEKVGREGVITVEDGSLENELDVNMKGQDFRGLSPYFINDAEKQIQAALDNP  
FVLLDFKRSINRDLPLVLEQVAKASRPILLIIEADVGEALATLVNNIRGLTKTVAV  
KAPGDFRRKAMLQDIAILITGGVWISSEVGLSEKATLDDLQGAQRKEITGKENTIID  
GFGDAAOIEARVAERQOETATSDYDEKLEQERAKLAGGVAVIKVGAATREVMKEK  
KDRVEDALHATRAAEVVEGVVAGGVALLPAALENLHTGNADQAGVQIVLRVESP  
LRQIVANAGSEPSVVNVKLEGNKNGYNAGSGEYGDMEIMGVLPDPAKVTLSALQHA  
STAGLMLTVDMAIEIPEDKPAVPMGGMGMGM"

BASE COUNT 2623 a 3083 c 2609 g 2276 t

Query Match 95.6%; Score 3062.2; DB 1; Length 10591;  
Best Local Similarity 97.1%; Pident. No. 0;  
Matches 3156; Conservative 0; Mismatches 48; Indels 45; Gaps 2;

QY 1 ATGCGAACGACCCCAACCTTCCCTACAAAACTTTTCAAAACGGCTGCCATGGCGTTAGCT 60  
|||||  
DB 548 ATGCGAACGACCCCAACCTTCCCTACAAAACTTTTCAAAACGGCTGCCATGGCGTTAGCT 607  
|||||  
QY 61 GTTGCAACAACTTCTTTCCTCTTAGCGGGCGGC---GGCGGCACCTTCTGCGCCCGAC 117  
|||||  
DB 608 GTTGCAACAACTTCTTTCCTCTTAGCGGGCGGCAGGCGGCACTTCTGCGCCCGAC 667  
|||||  
QY 118 TTCAATGCAGGCGGACCGGTATCGGCAGCAACAGCAGCAGCAACAGCGAATACAGCA 177  
|||||  
DB 668 TTCAATGCAGGCGGTATCGGCAGCAACAGCAGCAGCAACAGCGAATACAGCA 727  
|||||  
QY 178 GCAGTATCTTACCCCGGTATCAAGAAAGAAATGTGCAAGACAGCAAGAGTGTCTGTGCC 237  
|||||  
DB 728 GCAGTATCTTACCCCGGTATCAAGAAAGAAATGTGCAAGACAGCAAGAGTGTCTGTGCC 787  
|||||  
QY 238 GTTCGGGATGACGTTTCGGGTTACAGACAGGAGTGCACAAATCAATGCCCGCCCGAAT 297  
|||||  
DB 788 GTTCGGGATGACGTTTCGGGTTACAGACAGGAGTGCACAAATCAATGCCCGCCCGAAT 847  
|||||  
QY 298 CTGCATACCGGACACTTACAAACCCAAATGAGCAGTACAGAAATTTGATCAACCTCAAA 357  
|||||  
DB 848 CTGCATACCGGACACTTTCACAAACCCAAATGAGCAGTACAGAAATTTGATCAACCTCAAA 907  
|||||  
QY 358 CTGCAATTAAGCAGGCTATACAGGACGCGGGTAGAGTAGGTATCGTCGATACAGGC 417  
|||||  
DB 908 CTGCAATTAAGCAGGCTATACAGGACGCGGGTAGAGTAGGTATCGTCGATACAGGC 967  
|||||  
QY 418 GAATCCGTGGCAGCATATCTTTCCCGAACTGTATGGCAGAAAGAACACGCTATAAC 477  
|||||  
DB 968 GAATCCGTGGCAGCATATCTTTCCCGAACTGTATGGCAGAAAGAACACGCTATAAC 1027  
|||||  
QY 478 GAAAAATACAAAAACTATACGGCGGTATATCGGGAAGGAGCGCTGAAGACGAGGCGGT 537  
|||||  
DB 1028 GAAAAATACAAAAACTATACGGCGGTATATCGGGAAGGAGCGCTGAAGACGAGGCGGT 1087  
|||||  
QY 538 AAAGACATTAAGCTTCTTTCGACGATGAGGCGGTATAGAGACTGAAGCAAGACGCGACG 597  
|||||  
DB 1088 AAAGACATTAAGCTTCTTTCGACGATGAGGCGGTATAGAGACTGAAGCAAGACGCGACG 1147  
|||||  
QY 598 GATATCCGCCACGTAAGAAAGAAATCGGACACATCGATGTGGTCTCCCATATTTATTGGCGGG 657  
|||||  
DB 1148 GATATCCGCCACGTAAGAAAGAAATCGGACACATCGATTTGGTCTCCCATATTTATTGGCGGG 1207  
|||||  
QY 658 CGTTCGTGGACGGCAGACCTGCAGCGGTATTTCGCCCGCATCGCGCTACACATAATG 717  
|||||  
DB 1208 CGTTCGTGGACGGCAGACCTGCAGCGGTATTTCGCCCGCATCGCGCTACACATAATG 1267  
|||||  
QY 718 AATACCATGATGAACACCAAGAACAAATATCTCGACGCCATCGCATGATGGTC 777  
|||||  
DB 1268 AATACCATGATGAACACCAAGAACAAATATCTCGACGCCATCGCATGATGGTC 1327  
|||||  
QY 778 AAGCTGGGGAAGCTGGCGGTGCGCATCGTCAATAACAGATTTTGGAAACACATCGAGGCA 837  
|||||



Db 1328 AAGCTGGCGCAACGTGGCGTGGCGCATCGTCAATACAGTCTTTGGAAACAACATCGAGGGCA 1387  
Qy 838 GGCACCTGCGGACCAATTTCCAAATAGCCAAATTCGGAGGAGCAGTACCGCAAGCGTTGCTC 897  
Db 1388 GGCACCTGCGGACCAATTTCCAAATAGCCAAATTCGGAGGAGCAGTACCGCAAGCGTTGCTC 1447  
Qy 898 GCCTATTCCCGCGGTGATAAACAAGACAGAGGAGTATCCGCCGTGATGCAACAGAGCGATTAC 957  
Db 1448 GACTATTCCCGCGGTGATAAACAAGACAGAGGAGTATCCGCCGTGATGCAACAGAGCGATTAC 1507  
Qy 958 GGCACCTGCTCTACCAACATCCGTAATAAACAATGCTTTTCAATTTTTCGGCAGCAAT 1017  
Db 1508 GGCACCTGCTCTACCAACATCCGTAATAAACAATGCTTTTCAATTTTTCGGCAGCAAT 1567  
Qy 1018 GACGCACAAGCTCAGCCCAACACACTGACCTATTGCCATTTATGAAATAAGACTCTCAA 1077  
Db 1568 GACGCACAAGCTCAGCCCAACACACTATGCCCTATTGCCATTTATGAAATAAGACTCTCAA 1627  
Qy 1078 AAAGGCATTATCACAGTCCGAGGGGTAGACCGCATGCGGAGGAGGAGTTCA----- 1126  
Db 1628 AAAGGCATTATCACAGTCCGAGGGGTAGACCGCATGCGGAGGAGGAGTTCAAAACGGGAAATG 1687  
Qy 1127 -----ATGGCTCCAACCAATTCGGAATTACTGCC 1155  
Db 1688 TATGGAGAACCGGTTACAGAACCCGCTTGAGTATGGCTCCAACCAATTCGGAATTACTGCC 1747  
Qy 1156 ATGTGCTGCCATTATCGCACCCCTATGAAGCAAGCGTCCGTTTCACCCGTACAAACCCGATT 1215  
Db 1748 ATGTGCTGCCATTATCGCACCCCTATGAAGCAAGCGTCCGTTTCACCCGTACAAACCCGATT 1807  
Qy 1216 CAATTTGCGGAACATCTTTTCCGACCCCATGCGTAACCCGGACGGCGGCTCTGCTGCTG 1275  
Db 1808 CAATTTGCGGAACATCTTTTCCGACCCCATGCGTAACCCGGACGGCGGCTCTGCTGCTG 1867  
Qy 1276 CAGAAATACCGGTGATGAGCAACCAACCACTGCTACCAAGCTCTGACACAGCGCTCAG 1335  
Db 1868 CAGAAATACCGGTGATGAGCAACCAACCACTGCTACCAAGCTCTGACACAGCGCTCAG 1927  
Qy 1336 GACATCGGTGAGTCCGCGTGGACAGCAAGTTTCGGCTGGGACTGCTGGATCGCGGTAA 1395  
Db 1928 GACATCGGTGAGTCCGCGTGGACAGCAAGTTTCGGCTGGGACTGCTGGATCGCGGTAA 1987  
Qy 1396 GCCATGAACGGAACCGGCTCTTCCGTTCCGCGGACTTTACCGCGGATACGAAAGGTACA 1455  
Db 1988 GCCATGAACGGAACCGGCTCTTCCGTTCCGCGGACTTTACCGCGGATACGAAAGGTACA 2047  
Qy 1456 TCCGATATTGCTACTCTCTCCGTAACGACATTTTCAGGCACGGCGGCTGATCAAAAAA 1515  
Db 2048 TCCGATATTGCTACTCTCTCCGTAACGACATTTTCAGGCACGGCGGCTGATCAAAAAA 2107  
Qy 1516 GCGCGCAGCAACTGCAACTGCACGCGCAACACACTATACGGCAAAACCACTTATCGAA 1575  
Db 2108 GCGCGCAGCAACTGCAACTGCACGCGCAACACACTATACGGCAAAACCACTTATCGAA 2167  
Qy 1576 GCGGTTTCGCTGTTGTACGGCAACCAAAATCGGATATGCGGCTGCAAAACCAAGGT 1635  
Db 2168 GCGGTTTCGCTGTTGTACGGCAACCAAAATCGGATATGCGGCTGCAAAACCAAGGT 2227  
Qy 1636 GCGCTGATTTAATACGGGGCGCATCCGGCGTACGCTGAAACAGCGGCAATTCGCTAT 1695  
Db 2228 GCGCTGATTTAATACGGGGCGCATCCGGCGGCACTCGGCGGCACTGAAACAGCGGCAATTCGCTAT 2287  
Qy 1696 CTGCGCAGATACCGACGATCCGGCGCAACCAACCGCTGACATCAAAAGCGGATCTGCAG 1755  
Db 2288 CTGCGCAGATACCGACCAATCCGGCGCAACCAACCGCTGACATCAAAAGCGGATCTGCAG 2347  
Qy 1756 CTGCGCGGCAAGGTACGCTGTACACAGCTTTGGGCAACACTGCTGAAAGTGACCGGTACG 1815  
Db 2348 CTGCGCGCAAAAGGTACGCTGTACACAGCTTTGGGCAACACTGCTGAAAGTGACCGGTACG 2407  
Qy 1816 GCGATACCGCGGCAAGCTGTACATGTGCGCACCGGCAAGGGCGGCTATCTCAAC 1875  
Db 2408 GCGATATACCGCGGCAAGCTGTACATGTGCGCACCGGCAAGGGCGGCTATCTCAAC 2467

Qy 1876 CGTACCGGACAACGTTGTTCCCTTCTGAGTCCGCGCAAAATTCGGCGGGGATTTCTTTTC 1935  
Db 2468 AGTACCGGACGACGTTGTTCCCTTCTGAGTCCGCGCAAAATTCGGCGGGGATTTCTTTTC 2527  
Qy 1936 TTCAAAACATCGAAACGAGCGGTGCTGCTGGCTTCCCTCGACAGCGTGAAGAAACA 1995  
Db 2528 TTCAAAACATCGAAACGAGCGGTGCTGCTGGCTTCCCTCGACAGCGTGAAGAAACA 2587  
Qy 1996 GCGGCGAGTGAAGCGCACACGCTGCTATTATGTCCTCGCGCAATTCGCGCAGCGACT 2055  
Db 2588 GCGGCGAGTGAAGCGCACACGCTGCTATTATGTCCTCGCGCAATTCGCGCAGCGACT 2647  
Qy 2056 GCTTTCGCGAGCGCACATTCGCGCGCCCGCTCTGAACACGCGGTAGAAGAGGGCGC 2115  
Db 2648 GCTTTCGCGAGCGCACATTCGCGCGCCCGCTCTGAACACGCGGTAGAAGAGGGCGC 2707  
Qy 2116 AGCAATCTGGAACACCTGATGGTGAACCTGGATCCGCTCCGAATCATTCGCAACACCCGAG 2175  
Db 2708 AGCAATCTGGAACACCTGATGGTGAACCTGGATCCGCTCCGAATCATTCGCAACACCCGAG 2767  
Qy 2176 ACGGTTGAAACTGCGGCGCGCGACGACAGATATGCGCGGCATCCGCCCTACGCGCA 2235  
Db 2768 ACGGTTGAAACTGCGGCGCGCGACGACAGATATGCGCGGCATCCGCCCTACGCGCA 2827  
Qy 2236 ACTTTCGCGAGCGGACGCGTACAGCATGCGAATGCGCGCGGAGGTGTACGCAATCTTC 2295  
Db 2828 ACTTTCGCGAGCGGACGCGTACAGCATGCGAATGCGCGCGGAGGTGTACGCAATCTTC 2887  
Qy 2296 AACAGTCTGCGCGTACCGTCTATGCGCACAGTACCGCGCGCCCATGCGCGATGACAGGA 2355  
Db 2888 AACAGTCTGCGCGTACCGTCTATGCGCACAGTACCGCGCGCCCATGCGCGATGACAGGA 2947  
Qy 2356 CGCGGCTGAAAGCGTATCGGACGGTTGGACCAACAGCTACGGGTCTGCGCGTCAATC 2415  
Db 2948 CGCGGCTGAAAGCGTATCGGACGGTTGGACCAACAGCTACGGGTCTGCGCGTCAATC 3007  
Qy 2416 CGCAAAACCCCAACGAGCGGTGGAACAGTGGAAACAGGCGGTGTTGAAGGCAAAATGGCG 2475  
Db 3008 CGCAAAACCCCAACGAGCGGTGGAACAGTGGAAACAGGCGGTGTTGAAGGCAAAATGGCG 3067  
Qy 2476 GGCAGTACCCAAACCGTCCGCTATGCGCGGAAACCGCGGAAATACGACAGCGCGCC 2535  
Db 3068 GGCAGTACCCAAACCGTCCGCTATGCGCGGAAACCGCGGAAATACGACAGCGCGCC 3127  
Qy 2536 ACACGTGGCATGGGACACAGACATGGAGCGGAAACAGTGAACATGCAAAACCCGACAGC 2595  
Db 3128 ACACGTGGCATGGGACACAGACATGGAGCGGAAACAGTGAACATGCAAAACCCGACAGC 3187  
Qy 2596 ATTAGTCTGTTTTCAGGCAATACGGCACGATGCGGCGGATATCGGCTATCTCAAGGCGCTG 2655  
Db 3188 ATTAGTCTGTTTTCAGGCAATACGGCACGATGCGGCGGATATCGGCTATCTCAAGGCGCTG 3247  
Qy 2656 TTCTCTACGGACGCTACAAAACAGCATACCGCGACGCGGTGCGGAGCAACATCGG 2715  
Db 3248 TTCTCTACGGACGCTACAAAACAGCATACCGCGACGCGGTGCGGAGCAACATCGG 3307  
Qy 2716 GAAGCAGCGTCAACGCGCACGCTGATGACGTGGCGCACGTGGCGGTGTCAACGTTCCG 2775  
Db 3308 GAAGCAGCGTCAACGCGCACGCTGATGACGTGGCGCACGTGGCGGTGTCAACGTTCCG 3367  
Qy 2776 TTTGCCCAACGGGAGATTGACGGTTCGAAGCGGTCTGCGGTACGACCTGCTCAAAACAG 2835  
Db 3368 TTTGCCCAACGGGAGATTGACGGTTCGAAGCGGTCTGCGGTACGACCTGCTCAAAACAG 3427  
Qy 2836 GATGATTCGCGGAAAAAGGAGTCTTTGGCTGGAGGGCAACAGCTCTACTGAAGGC 2895  
Db 3428 GATGATTCGCGGAAAAAGGAGTCTTTGGCTGGAGGGCAACAGCTCTACTGAAGGC 3487  
Qy 2896 ACACGTGCGGACTCGCGGCTGAAGCTGTCGCAACCTTTCAGCGATAAAGCCGCTCTG 2955  
Db 3488 ACCTGCTGCGNCTCGCGGCTGAAGCTGTCGCAACCTTTCAGCGATAAAGCCGCTCTG 3547



Qy	2956	TTTGCACAGCGCGCGGTGGAAACGCGACCTGAACAGGACGCGACTACACGGTAACGGGGCGG	3015
Db	3548	TTTGCACAGCGCGCGGTGGAAACGCGACCTGAACAGGACGCGACTACACGGTAACGGGGCGG	3607
Qy	3016	TTTACCGCGCGGACTGCGAGCAACCGGCAAGACGGGGGCGACGCAATATGCCGCACACCCG	3075
Db	3508	TTTACCGCGCGGACTGCGAGCAACCGGCAAGACGGGGGCGACGCAATATGCCGCACACCCG	3667
Qy	3076	CTGGTTCCCGGCTGTGGCGCGGGATGTGCAATTTCGGCAACGGCTGGAACGGCTTGGCACGT	3135
Db	3668	CTGGTTCCCGGCTGTGGCGCGGGATGTGCAATTTCGGCAACGGCTGGAACGGCTTGGCACGT	3727
Qy	3136	TACAGCTACGCGGTTTCCAAACAGTACGGCAACACAGCGGACGAGTGGCGGTAGGCTAC	3195
Db	3728	TACAGCTACGCGGTTTCCAAACAGTACGGCAACACAGCGGACGAGTGGCGGTAGGCTAC	3787
Qy	3196	CGGTTCTGA 3204	
Db	3788	CGGTTCTGA 3796	
RESULT 5			
AX044034			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
BASE COUNT			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 3156; Conservative			
Qy 1			
Db 261412			
Qy 61			
Db 261472			
Qy 118			
Db 261532			
Qy 178			
Db 261592			

Qy	238	GGTCCGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAATGCCCCCCCCCCGAAT	297
Db	261652	GGTCCGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAATGCCCCCCCCCCGAAT	261711
Qy	298	CTGCATACCGGAGACATTACAAACCCAAATGACGCATACAGAATTTGATCAACCTCAA	357
Db	261712	CTGCATACCGGAGACTTCCAAACCCAAATGACGCATACAGAATTTGATCAACCTCAA	261771
Qy	358	CCTGCAATTGAAGCAGGCTATACGAGACGCGGGTAGAGTAGGTAGGTATCGTCGATACAGGC	417
Db	261772	CCTGCAATTGAAGCAGGCTATACGAGACGCGGGTAGAGTAGGTAGGTATCGTCGATACAGGC	261831
Qy	418	GAATCCGTCGGCAGCATATCCTTTCCGGAACGTATGGCAGAAAAAGAACACGGCTATAAC	477
Db	261832	GAATCCGTCGGCAGCATATCCTTTCCGGAACGTATGGCAGAAAAAGAACACGGCTATAAC	261891
Qy	478	GAAATTTACAAAACATATACGGCGTATATCGGGAAGAACGCGCTGAAGACGGAGCGGT	537
Db	261892	GAAATTTACAAAACATATACGGCGTATATCGGGAAGAACGCGCTGAAGACGGAGCGGT	261951
Qy	538	AAAGACATTAAGGCTCTTTTCGACGATGAGCGCGTTATAGAGACTGAAGCAAAAGCCGACG	597
Db	261952	AAAGACATTAAGGCTCTTTTCGACGATGAGCGCGTTATAGAGACTGAAGCAAAAGCCGACG	262011
Qy	598	GATATCCGCCACGTAAGAAAAATCGGACATATCGATGTGGTCTCCCATATTTATGGCGGG	657
Db	262012	GATATCCGCCACGTAAGAAAAATCGGACATATCGATGTGGTCTCCCATATTTATGGCGGG	262071
Qy	658	CGTTCCGTGGACGGCAGACCTGACGGGGTATTTGCCCGCGATGCGACGCTACACATAATG	717
Db	262072	CGTTCCGTGGACGGCAGACCTGACGGGGTATTTGCCCGCGATGCGACGCTACACATAATG	262131
Qy	718	AATACGCATGATGAACCAAGAACAAATAATCTCTGACGCCATCGCAATGCATCGTGC	777
Db	262132	AATACGAATGATGAACCAAGAACAAATGATGGTTGCAAGCCATCGCAATGCATCGTGC	262191
Qy	778	AAGTGGCGGAACGTGGCGTCGCATCGTCAATAACAGTTTTTGGAAACACATCGAGGGCA	837
Db	262192	AAGTGGCGGAACGTGGCGTCGCATCGTCAATAACAGTTTTTGGAAACACATCGAGGGCA	262251
Qy	838	GGCACTCCGACACATTTCCAAATAGCCATTTCCGAGGACGATACCGCCAAAGCTTGCTC	897
Db	262252	GGCACTCCGACACATTTCCAAATAGCCATTTCCGAGGACGATACCGCCAAAGCTTGCTC	262311
Qy	898	GCGTATTTCCGGCGGTGATAAACAAGACGAGGGTATCCGCTGATGCAACAGACGATTAC	957
Db	262312	GACTATTTCCGGCGGTGATAAACAAGACGAGGGTATCCGCTGATGCAACAGACGATTAC	262371
Qy	958	GGCAACTTGTCCTACACATCCGTAATAAACAATGCTTTTCAATTTTTCGGCAAGCAAT	1017
Db	262372	GGCAACCTGTCCTACACATCCGTAATAAACAATGCTTTTCAATTTTTCGGCAAGCAAT	262431
Qy	1018	GACGCAAGCTCAGCCCAACACACTGACCCCTATTCGCATTTTATGAAAAAGATCCTCAA	1077
Db	262432	GACGCAAGCTCAGCCCAACACATATGGCCCTATTCGCATTTTATGAAAAAGATCCTCAA	262491
Qy	1078	AAAGGCATTTACAGTCGACGGGTAGACCGCAGTGGAGAAAAGTTCA-----1126	1126
Db	262492	AAAGGCATTTACAGTCGACGGGTAGACCGCAGTGGAGAAAAGTTCAACCGGAATG	262551
Qy	1127	-----ATGGCTCCAACCATTTGGGGAATTTACTGCC1155	1155
Db	262552	TATGGAGAACCGGTACAGAACCGCTTGAGTATGGCTCCAACCATTTGGGGAATTTACTGCC	262611
Qy	1156	ATGTTGTTGCCTATCGGCACCCCTATGAAGCAAGGTCGTTTCAACCGTACAACCCGATT	1215
Db	262612	ATGTTGTTGCCTATCGGCACCCCTATGAAGCAAGGTCGTTTCAACCGTACAACCCGATT	262671
Qy	1216	CAAAATCCCGGAACATCCTTTTCCGACCCCATCGTAACCGGACGCGGCTCTGCTGCTG	1275
Db	262672	CAAAATCCCGGAACATCCTTTTCCGACCCCATCGTAACCGGACGCGGCTCTGCTGCTG	262731
Qy	1276	CAGAAATACCCGTGGATTGAGCAACGACACTCGCTACCAACGCTGCTGCAACAGGCTCAG	1335

|||||  
Db 262732 CAGAAATACCGTGTAGCAGCAACACACCTGCGTACACGTTCTGACGACGCTCAG 262791  
QY 1336 GACATCGGTGAGTCGGGTGGACAGAGTTTCGGCTGGGACTCTGATCGGGTAAG 1395  
Db 262792 GACATCGGTGAGTCGGGTGGACAGAGTTTCGGCTGGGACTCTGATCGGGTAAG 262851  
QY 1396 GCCATGAACGACCGCGCTCTTTCCGTTTCGGCGACTTTACCGCGGATACGAAAGGTACA 1455  
Db 262852 GCCATGAACGACCGCGCTCTTTCCGTTTCGGCGACTTTACCGCGGATACGAAAGGTACA 262911  
QY 1456 TCCGATATTGCTACTCTTCGTTAACGACATTTTCAGGACGAGGGCGCTGATCAAAAA 1515  
Db 262912 TCCGATATTGCTACTCTTCGTTAACGACATTTTCAGGACGAGGGCGCTGATCAAAAA 262971  
QY 1516 GCGCGACGCAACTGCAACTGCACGGCAACACACCTATACGGGCAAAACCATTTATCGAA 1575  
Db 262972 GCGCGACGCAACTGCAACTGCACGGCAACACACCTATACGGGCAAAACCATTTATCGAA 263031  
QY 1576 GCGGTTTCGCTGGTGTGTAGGCAACAACTCGGATATGCGGTCGAAACCAAGGT 1635  
Db 263032 GCGGTTTCGCTGGTGTGTAGGCAACAACTCGGATATGCGGTCGAAACCAAGGT 263091  
QY 1636 GCGCTGATTTAAGCGGGCGGATCCGGCGGTAGCCTGACAGCGAGGATTTGCTAT 1695  
Db 263092 GCGCTGATTTAAGCGGGCGGATCCGGCGGTAGCCTGACAGCGAGGATTTGCTAT 263151  
QY 1696 GTGGCAGATACCGACCGATCCGGCGCAACAAACCGTGCACATCAAGGGCATCTGCAG 1755  
Db 263152 GTGGCAGATACCGACCGATCCGGCGCAACAAACCGTGCACATCAAGGGCATCTGCAG 263211  
QY 1756 CTGGCGCGGAGGTACGCTGTACACAGCTTTGGGCAAACTGCTGAAAGTGACGCTACG 1815  
Db 263212 CTGGCGCGGAGGTACGCTGTACACAGCTTTGGGCAAACTGCTGAAAGTGACGCTACG 263271  
QY 1816 GCGATACCGCGGCAAGCTGTACATGTGGCACCGCGCAAGAGGGAGGATATCTCAAC 1875  
Db 263272 GCGATTTATCGCGGCAAGCTGTACATGTGGCACCGCGCAAGAGGGAGGATATCTCAAC 263331  
QY 1876 GGTACCGGACAGGTGTTCCCTCTCTAGTGCAGCGCAAAATCGGGCGGATTTCTTTTC 1935  
Db 263332 AGTACCGGACAGGTGTTCCCTCTCTAGTGCAGCGCAAAATCGGGCGGATTTCTTTTC 263391  
QY 1936 TTCACAAACATCGAAACGACGCTGTCTGCTGCTCCCTCGACAGCGTCGAAAAACA 1995  
Db 263392 TTCACAAACATCGAAACGACGCTGTCTGCTGCTCCCTCGACAGCGTCGAAAAACA 263451  
QY 1996 GCGGCAGTGAAGGCGACAGCTGTCTATTATGTCTGGGGCAATGCGGCAGGACT 2055  
Db 263452 GCGGCAGTGAAGGCGACAGCTGTCTATTATGTCTGGGGCAATGCGGCAGGACT 263511  
QY 2056 GCTTCGGCAGCGGCACATTCGCGCGCCGCTCTGAAACACGCGGTAGAACAGGGCGGC 2115  
Db 263512 GCTTCGGCAGCGGCACATTCGCGCGCCGCTCTGAAACACGCGGTAGAACAGGGCGGC 263571  
QY 2116 AGCAATCTGGAACCTGATGGTGAATCGGATGCTCCGAAATCATCCGCAACACCCGAG 2175  
Db 263572 AGCAATCTGGAACCTGATGGTGAATCGGATGCTCCGAAATCATCCGCAACACCCGAG 263631  
QY 2176 ACGGTTGAATCGGGCGCGCAGCGCACACATATGCGGGGATCCGCGCCCTACGGCGCA 2235  
Db 263632 ACGGTTGAATCGGGCGCGCAGCGCACACATATGCGGGGATCCGCGCCCTACGGCGCA 263691  
QY 2236 ACTTTCGCGCAGCGGACGCTACAGCATCGGAATCGCGGACGGGTGACGATCTTC 2295  
Db 263692 ACTTTCGCGCAGCGGACGCTACAGCATCGGAATCGCGGACGGGTGACGATCTTC 263751  
QY 2296 AACAGTCTCGCGCTACCGTCTATGCCGACAGTACCGCGCCCATCGCGATATCGAGGA 2355  
Db 263752 AACAGTCTCGCGCTACCGTCTATGCCGACAGTACCGCGCCCATCGCGATATCGAGGA 263811  
QY 2356 GCGCGGCTGAAGCGGTATCGGACGGGTGAGCCACACGCTACGGGTCGCGGCTCATC 2415  
|||||

Db 263812 CGCGCCCTGAAGCGGTATCGGACGCGTTGGACCAACAGGCACGCGTCTGCGGCTCATC 263871  
QY 2416 GCGCAAAACCCAAACAGGACGCTGGAAACGTTGGAAACAGGCGGTGTTGAAGGAAAAATCGC 2475  
Db 263872 GCGCAAAACCCAAACAGGACGCTGGAAACGTTGGAAACAGGCGGTGTTGAAGGAAAAATCGC 263931  
QY 2476 GCGAGTACCCAAACCGTGGGATTCGCCGAAACCCGCGGAAAAATACGACAGCAGCCGC 2535  
Db 263932 GCGAGTACCCAAACCGTGGGATTCGCCGAAACCCGCGGAAAAATACGACAGCAGCCGC 263991  
QY 2536 ACAGTGGCATGGGACACAGCAGCATGGAGCGAAACAGTGCAGAAATGCAAAAACCCAGC 2595  
Db 263992 ACAGTGGCATGGGACGACGACATGGAGCGAAACAGTGCAGAAATGCAAAAACCCAGC 264051  
QY 2596 ATTAGTCTGTTTGGAGGATACGGCAGCATGGGCGGATATCGGCTATCTCAAGGCGCTG 2655  
Db 264052 ATTAGTCTGTTTGGAGGATACGGCAGCATGGGCGGATATCGGCTATCTCAAGGCGCTG 264111  
QY 2656 TTCTCCTACGGACGCTACAAAAACAGCATCAGCCGACGACCGGTGCGGAGCAACATCGC 2715  
Db 264112 TTCTCCTACGGACGCTACAAAAACAGCATCAGCCGACGACCGGTGCGGAGCAACATCGC 264171  
QY 2716 GAAGCAGCGTCAACGGCACCGTGTATGACGCTGGCGGACCTGGGCGGTGTCAACGTTCCG 2775  
Db 264172 GAAGCAGCGTCAACGGCACCGTGTATGACGCTGGCGGACCTGGGCGGTGTCAACGTTCCG 264231  
QY 2776 TTTGCCGCAACGGGAGATTTGACGCTCGAAGCGGTCTGCGCTACGACCTGCTCAACAG 2835  
Db 264232 TTTGCCGCAACGGGAGATTTGACGCTCGAAGCGGTCTGCGCTACGACCTGCTCAACAG 264291  
QY 2836 GATGCATTCGCGGAAAAAGGAGTGTCTTTGGGCTGGAGCGGCAACAGCCTCACTGAAGGC 2895  
Db 264292 GATGCATTCGCGGAAAAAGGAGTGTCTTTGGGCTGGAGCGGCAACAGCCTCACTGAAGGC 264351  
QY 2896 ACAGTGTGCGACTCGCGGCTGTGAAGCTGTGCAACCTTTGAGGCGATAAGCCGCTCG 2955  
Db 264352 ACGTGTGCGACTCGCGGCTGTGAAGCTGTGCAACCTTTGAGGCGATAAGCCGCTCG 264411  
QY 2956 TTTGCAACGGCGGCGTGGAAACGGGACCTGAACGACGCGACTACACGCTAACGGCGGC 3015  
Db 264412 TTTGCAACGGCGGCGTGGAAACGGGACCTGAACGACGCGACTACACGCTAACGGCGGC 264471  
QY 3016 TTTACCGCGCGACTGTGACGCAACCGGCAAGAGCGGGGACGCAATATCCGCGACACCCGC 3075  
Db 264472 TTTACCGCGCGACTGTGACGCAACCGGCAAGAGCGGGGACGCAATATCCGCGACACCCGC 264531  
QY 3076 CTGGTTCGGCTCTGGGCGCGGATGTCGAATTCGCAACGCGTGGAAAGGCTTGGCAGCT 3135  
Db 264532 CTGGTTCGGCTCTGGGCGCGGATGTCGAATTCGCAACGCGTGGAAAGGCTTGGCAGCT 264591  
QY 3136 TACAGCTACGCGGTTCCAAACAGTACGCAACACAGCGGACGAGTGGGCTAGGCTAC 3195  
Db 264592 TACAGCTACGCGGTTCCAAACAGTACGCAACACAGCGGACGAGTGGGCTAGGCTAC 264651  
QY 3196 CGGTTCTGA 3204  
|||||  
Db 264652 CGGTTCTGA 264660

RESULT 6  
NME277537  
LOCUS  
DEFINITION NME277537 3254 bp DNA linear BCT 24-JUL-2002  
Neisseria meningitidis ausp gene for autotransporter serine  
protease.  
ACCESSION AJ277537  
VERSION AJ277537.2 GI:13539237  
KEYWORDS ausp gene; autotransporter serine protease.  
SOURCE Neisseria meningitidis.  
ORGANISM Neisseria meningitidis  
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;  
Neisseria.  
REFERENCE 1  
AUTHORS Turner,D.P., Wooldridge,K.G. and Ala'Aldeen,D.A.

TITLE Autotransported Serine Protease A of *Neisseria meningitidis*: an  
JOURNAL Immunogenic, Surface-Exposed Outer Membrane, and Secreted Protein  
MEDLINE Infect. Immun. 70 (8), 4447-4461 (2002)  
REFERENCE 22112893  
AUTHORS Ala'Aldeen, D.A.A.  
TITLE Direct Submission  
JOURNAL Submitted (19-APR-2000) Ala'Aldeen D.A.A., Microbiology, University  
of Nottingham, Meningococcal Research Group, Division of  
Microbiology, University Hospital, Nottingham NG7 2UH, UNITED  
KINGDOM  
REMARK Revised by [3]  
REFERENCE 3 (bases 1 to 3254)  
AUTHORS Ala'Aldeen, D.A.A.  
TITLE Direct Submission  
JOURNAL Submitted (02-APR-2001) Ala'Aldeen D.A.A., Microbiology, University  
of Nottingham, Meningococcal Research Group, Division of  
Microbiology, University Hospital, Nottingham NG7 2UH, UNITED  
KINGDOM  
COMMENT On Apr 3, 2001 this sequence version replaced gi:7649689.  
FEATURES  
source Location/Qualifiers  
1..3254  
/organism="Neisseria meningitidis"  
/strain="B:15:P1.16"  
/isolate="SD"  
/db\_xref="taxon:487"  
/country="United Kingdom:England"  
1..3254  
/gene="ausp"  
/pseudo  
1..3254  
/gene="ausp"  
/function="unknown"  
/pseudo  
/codon\_start=1  
/trans\_table=11  
/product="autotransporter serine protease"  
BASE COUNT 845 a 922 c 917 g 570 t

Query Match 95.0%; Score 3042.4; DB 1; Length 3254;  
Best Local Similarity 96.9%; Pred. No. 0;  
Matches 3153; Conservative 0; Mismatches 51; Indels 50; Gaps 3;  
QY 1 ATGCGAACGACCCCACTTCCTACAAAACTTTCAAAACGGCTGCCATGGCGTTAGCT 60  
DB 1 ATGCGAACGACCCCACTTCCTACAAAACTTTCAAAACGGCTGCCATGGCGTTAGCT 60  
QY 61 GTTGCAACAACACTTCTGCGCTGTAGCGCGCGC---GGCGGCACTTCTGCGCCGAC 117  
DB 61 GTTGCAACAACACTTCTGCGCTGTAGCGCGCGGAGGCGGCACCTTCTGCGCCGAC 120  
QY 118 TTCATGCGGGGACCGGATATCGGAGCACACAGCAGCAGCAACAGCGAATCAGCA 177  
DB 121 TTCATGCGGGGATACCGGATATCGGAGCACACAGCAGCAGCAACAGCGAATCAGCA 180  
QY 178 GCAGTATCTTACGCGGGTATCAAGAACAAATGTGCAAGACAGAGCATGCTGTGTC 237  
DB 181 GCAGTATCTTACGCGGGTATCAAGAACAAATGTGCAAGACAGAGCATGCTGTGTC 240  
QY 238 GGTGGGATGAGTGTGCGGTTACAGACAGGGATGCCAAATCAATG-----CCCCCCCC 292  
DB 241 GGTGGGATGAGTGTGCGGTTACAGACAGGGATGCCAAATCAATGCCCCCCCCCCCC 300  
QY 293 CGAATCTGCATACCGGACACTTTCAAAACCCAAATGACGCATACAAAGATTTGATCAAC 352  
DB 301 CGAATCTGCATACCGGACACTTTCAAAACCCAAATGACGCATACAAAGATTTGATCAAC 360  
QY 353 TCAAACTTGCATTAAGCAGGCTATACAGACGGGGGTAGAGTAGGTATCGTCGATA 412  
DB 361 TCAAACTTGCATTAAGCAGGCTATACAGACGGGGGTAGAGTAGGTATCGTCGATA 420  
QY 413 CAGCGAATCCGTGGGACGATATCCTTTCCGGAATGTATGGCAGAAAAAGAACGCGCT 472

DB 421 CAGGCGAATCCGTGGGACGATATCCTTTCCGGAATGTATGGCAGAAAAAGAACGCGCT 480  
QY 473 ATAACGAAAAATTACAAAACTATACCGCGTATATCGGAGGAGGCGCTGAAGACGGAG 532  
DB 481 ATAACGAAAAATTACAAAACTATACCGCGTATATCGGAGGAGGCGCTGAAGACGGAG 540  
QY 533 GCGGTAAGACATTAAGCTTCTTCGAGGATGAGCGCGTTATAGAGACTGAAGCAAGC 592  
DB 541 GCGGTAAGACATTAAGCTTCTTCGAGGATGAGCGCGTTATAGAGACTGAAGCAAGC 600  
QY 593 CGACGGATATCCGCCACGTAAAGAAATCGGACATCGATGTGTCTCCCATATTATTG 652  
DB 601 CGACGGATATCCGCCACGTAAAGAAATCGGACATCGATGTGTCTCCCATATTATTG 660  
QY 653 GCGGCGTTCCTGGGACGACAGCTGCGGCGTATTCGCCCGGATGCGGCGCTACACA 712  
DB 661 GCGGCGTTCCTGGGACGACAGCTGCGGCGTATTCGCCCGGATGCGGCGCTACACA 720  
QY 713 TAATGAATACGATGATGGAACAAGCAATAATGTCTGCAGGCATCCGCAATGCAT 772  
DB 721 TAATGAATACGATGATGGAACAAGCAATAATGTGTGCAGGCATCCGCAATGCAT 780  
QY 773 GGGTCAAGCTGGGCAACGTGGCGTCGCATCGTCAATAACAGTTTGGAAACAATCGA 832  
DB 781 GGGTCAAGCTGGGCAACGTGGCGTCGCATCGTCAATAACAGTTTGGAAACAATCGA 840  
QY 833 GGGCAGGCACTCCGACCAATTTCCAAATAGCCAAATTCGGAGGAGCAGTACCGCCAGCGT 892  
DB 841 GGGCAGGCACTCCGACCAATTTCCAAATAGCCAAATTCGGAGGAGCAGTACCGCCAGCGT 900  
QY 893 TGCTCGCTATTTCGGCGGTGATAAAACAGAGAGGTATCCGCTGATGCAACAGAGCG 952  
DB 901 TGCTCGACTATTTCGGCGGTGATAAAACAGAGAGGTATCCGCTGATGCAACAGAGCG 960  
QY 953 ATTACGGCAACTTGCTTACCACATCCGTAATAAAACATGCTTTTCATTTTCGCGCAA 1012  
DB 961 ATTACGGCAACTTGCTTACCACATCCGTAATAAAACATGCTTTTCATCTTTTCGACAG 1020  
QY 1013 GCAATGCGCACAGCTCAGCCCAACACACTGACCTTATGCAATTTTATGAAAAAGATG 1072  
DB 1021 GCAATGCGCACAGCTCAGCCCAACACATATGCCCTATTTGCAATTTATGAAAAAGAGC 1080  
QY 1073 CTCAAAAAGGCATTTATCAGTTCGCGGCTAGACCGCAGTGGAGAAAGTTCA----- 1126  
DB 1081 CTCAAAAAGGCATTTATCAGTTCGCGGCTAGACCGCAGCGAGAGAAAGTTCAACACGGG 1140  
QY 1127 -----ATGGCTCCCAACCACTTTCGGAATTA 1150  
DB 1141 AAATGTATGGAGAACCGGATACAGAACCGCTTGTAGTATGCTTCCCAACCACTTTCGGAATTA 1200  
QY 1151 CTGCGCATGTGTCCTATCGGCACCGCTATGAAGCAAGCGTCCGTTTACCCCGTACAAAC 1210  
DB 1201 CTGCGCATGTGTCCTATCGGCACCGCTATGAAGCAAGCGTCCGTTTACCCCGTACAAAC 1260  
QY 1211 CGATTCAAAATTCGCGGACATCTTTTCGCGACCCATCGTAACCGCAGCGCGCTCTGC 1270  
DB 1261 CGATTCAAAATTCGCGGACATCTTTTCGCGACCCATCGTAACCGCAGCGCGCTCTGC 1320  
QY 1271 TGCTGCAAAATPACCGGTGATGAGCAAGCAACCTGCGTACCACGCTGCTGCAACACGG 1330  
DB 1321 TGCTGCAAAATPACCGGTGATGAGCAAGCAACCTGCGTACCACGCTGCTGCAACACGG 1380  
QY 1331 CTCAGGACATCGGTGATGCGGCTGGGACAGCAAGTTTCGGCTGGGAGTGGGATGCGG 1390  
DB 1381 CTCAGGACATCGGTGATGCGGCTGGGACAGCAAGTTTCGGCTGGGAGTGGGATGCGG 1440  
QY 1391 GTAAGCCATGAACGACCGCGCTCTTTCCGCTCGGCGACTTTACCGCGGATGCAAGAG 1450  
DB 1441 GTAAGCCCTTGAACGACCGCGCTCTTTCCGCTCGGCGACTTTACCGCGGATGCAAGAG 1500  
QY 1451 GTACATCCGATTTGCTTACTCTTCCGTAACGACATTTTCAGGACGCGGCGCTGTATCA 1510





|||||  
Db 3216 TATGCGGCGATCCGCCCTACGGCGCAACTTTCGCGCAGCGGCGGCGGTACAGCATGC 3275  
Qy 2268 GAATCCGCCGACGGGTGTACGCATCTTCAACAGTCTCGCGGTACCGTCTATGCGGCAGC 2327  
Db 3276 GAATCCGCCGACGGGTGTACGCATCTTCAACAGTCTCGCGGTACCGTCTATGCGGCAGC 3335  
Qy 2328 TACCGCCGCCATGCGGATATGACAGGACGCGCGGTGAAAGCGGTATCGGACGGGTGGA 2387  
Db 3336 TACCGCCGCCATGCGGATATGACAGGACGCGCGGTGAAAGCGGTATCGGACGGGTGGA 3395  
Qy 2388 CCACAAAGCTAGGGTCTGCGGTCTATCGCGCAACCAACAGAGCGGTGGAACGTGGA 2447  
Db 3396 CCACAAAGCTAGGGTCTGCGGTCTATCGCGCAACCAACAGAGCGGTGGAACGTGGA 3455  
Qy 2448 ACAGGGCGGTGTGAAGCAAAATGCGGCGAGTACCCAAACCCGTCGCAATGCGCGCAA 2507  
Db 3456 ACAGGGCGGTGTGAAGCAAAATGCGGCGAGTACCCAAACCCGTCGCAATGCGCGCAA 3515  
Qy 2508 AACCGCGCAAAATACAGACAGCGCCACACTGGGCGATGGACACAGCATGAGGGA 2567  
Db 3516 AACCGCGCAAAATACAGACAGCGCCACACTGGGCGATGGACACAGCATGAGGGA 3575  
Qy 2568 AAACAGTGCAAATGCAAAACCGACAGCATTAGTCTGTTGAGGCGATACGGCAGCATGC 2627  
Db 3576 AAACAGTGCAAATGCAAAACCGACAGCATTAGTCTGTTGAGGCGATACGGCAGCATGC 3635  
Qy 2628 GGGCGATATCGGTATCTCAAGGCGCTGCTCTACCGAGCGCTCAAAACAGCATCAG 2687  
Db 3636 GGGCGATATCGGTATCTCAAGGCGCTGCTCTACCGAGCGCTCAAAACAGCATCAG 3695  
Qy 2688 CCGCAGCACCGTGGGAGCAATCGGAGCGAGCGCTCAACGGCAGCGTATGACGCT 2747  
Db 3696 CCGCAGCACCGTGGGAGCAATCGGAGCGAGCGCTCAACGGCAGCGTATGACGCT 3755  
Qy 2748 GGGCGCATGGGCGGTGTCAAGCTTCGTTGCGCGAAGCGGAGATTTGACGGTGCAGG 2807  
Db 3756 GGGCGCATGGGCGGTGTCAAGCTTCGTTGCGCGAAGCGGAGATTTGACGGTGCAGG 3815  
Qy 2808 CGGTCTGGCTACGACCTGCTCAACAGAGTCAATTCGCGGAAAGGAGTCTTTGG 2867  
Db 3816 CGGTCTGGCTACGACCTGCTCAACAGAGTCAATTCGCGGAAAGGAGTCTTTGG 3875  
Qy 2868 CTGGAGCGCAACAGCCTCACTGAAGGCACACTGGTGGGACTCGCGGTCTGAAGCTGTC 2927  
Db 3876 CTGGAGCGCAACAGCCTCACTGAAGGCAGCGTGTGGGACTCGCGGTCTGAAGCTGTC 3935  
Qy 2928 GCAACCCCTTGACGATGAAGCGCTCTGTTGCAACGGGCGGCGTGGACGGACCTGAA 2987  
Db 3936 GCAACCCCTTGACGATGAAGCGCTCTGTTGCAACGGGCGGCGTGGACGGACCTGAA 3995  
Qy 2988 CGGACCGGACTACAGGTAAAGCGGCGCTTTACCGGCGGCGACTGCAACCGGCAAGAC 3047  
Db 3996 CGGACCGGACTACAGGTAAAGCGGCGCTTTACCGGCGGCGACTGCAACCGGCAAGAC 4055  
Qy 3048 GGGGCGACGCAATATCGCGCACACCGCGCTGTTGCGCGTCTGGCGGGGATGTCGAATT 3107  
Db 4056 GGGGCGACGCAATATCGCGCACACCGCGCTGTTGCGCGTCTGGCGGGGATGTCGAATT 4115  
Qy 3108 CGGCAACGGCTGGAACGGTTGGACAGTTTACAGCTACGCGCGGTTCGAACAGTACGGCAA 3167  
Db 4116 CGGCAACGGCTGGAACGGTTGGACAGTTTACAGCTACGCGCGGTTCGAACAGTACGGCAA 4175  
Qy 3168 CCACAGCGGACAGTTCGGCGTAGGCTACCGGTTCTGA 3204  
Db 4176 CCACAGCGGACAGTTCGGCGTAGGCTACCGGTTCTGA 4212

RESULT 8  
LOCUS AX239755 4218 bp DNA linear PAT 26-SEP-2001  
DEFINITION Sequence 57 from Patent WO0164920.  
ACCESSION AX239755

AX239755.1 GI:15797388  
synthetic construct.  
synthetic construct  
artificial sequences.  
1 (bases 1 to 4218)  
Arico,M.B., Comanducci,M., Galeotti,C., Masignani,V., Giuliani,M.M.  
and Pizzi,M.  
Hybrid expression of neisserial proteins  
Patent: WO 0164920-A 57 07-SEP-2001;  
Chiron Spa (IT)  
Location/Qualifiers  
1..4218  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="961cL-983"  
BASE COUNT 1196 a 1149 c 1130 g 743 t  
ORIGIN  
Query Match 93.2%; Score 2986.2; DB 6; Length 4218;  
Best Local Similarity 97.1%; Pred. No. 0;  
Matches 3067; Conservative 0; Mismatches 48; Indels 42; Gaps 1;  
Qy 90 CGCGGCGCGGCACTTCTGCGCGGCACTTCAATGACGGCGGCAACCGGTATCGGCAGCAA 149  
Db 1056 CGCGGAGCGGCACTTCTGCGCGGCACTTCAATGACGGCGGTACCGGTATCGGCAGCAA 1115  
Qy 150 CAGCAGACCAACACAGCAATCAGCAGCAGTCTTACGCCGTATCAAGNACGAAT 209  
Db 1116 CAGCAGACCAACACAGCAATCAGCAGCAGTCTTACGCCGTATCAAGNACGAAT 1175  
Qy 210 GTGCAAAACAGACAGCAGTCTGTCGCCGTCGGGATGACGTTGCGGTTACAGACAGGGA 269  
Db 1176 GTGCAAAACAGACAGCAGTCTGTCGCCGTCGGGATGACGTTGCGGTTACAGACAGGGA 1235  
Qy 270 TGC AAAATCAATGCCCGCCCGGAAATCTGATACCGGAGACTTTTACAAACCCAAATGA 329  
Db 1236 TGCCAAAATCAATGCCCGCCCGGAAATCTGATACCGGAGACTTTTCCAAAACCCAAATGA 1295  
Qy 330 CGCATACAGAAATTTGATCAACCTCAACCTTGAAGCAGGCTATACAGGACGCG 389  
Db 1296 CGCATACAGAAATTTGATCAACCTCAACCTTGAAGCAGGCTATACAGGACGCG 1355  
Qy 390 GGTAGAGTAGGTATGTCGATACAGGCGAAATTCGTCGGCAGCATATCCTTTCCCGAACT 449  
Db 1356 GGTAGAGTAGGTATGTCGATACAGGCGAAATTCGTCGGCAGCATATCCTTTCCCGAACT 1415  
Qy 450 GTATGGCAGAAAGAACACGCGTATACGAAAATTTACAAAACATATACGGGTATATGCG 509  
Db 1416 GTATGGCAGAAAGAACACGCGTATACGAAAATTTACAAAACATATACGGGTATATGCG 1475  
Qy 510 GAAGGAAGCGCTGAAGACGAGGCGGTAAAGACATTTAAAGCTTCTTTCGACGATGAGGC 569  
Db 1476 GAAGGAAGCGCTGAAGACGAGGCGGTAAAGACATTTAAAGCTTCTTTCGACGATGAGGC 1535  
Qy 570 CGTTATAGAGACTGAAGCAAGCCGACGATATCGCCACGCTAAAGAAATCGGACACAT 629  
Db 1536 CGTTATAGAGACTGAAGCAAGCCGACGATATCGCCACGCTAAAGAAATCGGACACAT 1595  
Qy 630 CGATGTGCTCTCCCATATTTTGGCGGGGCTTCCGTGACGCGCAGACCTGACGGCGGTAT 689  
Db 1596 CGATTTGCTCTCCCATATTTTGGCGGGGCTTCCGTGACGCGCAGACCTGACGGCGGTAT 1655  
Qy 690 TGCGCCGATCGGCGCTACACATTAATGAATACGATATGGAACCAAGCAAGCAATAT 749  
Db 1656 TGCGCCGATCGGCGCTACACATTAATGAATACGATATGGAACCAAGCAAGCAATAT 1715  
Qy 750 GTCTGACGCCATCCGCAATGCGTCAAGCTGGCGGAAAGCTGCGGTGCGCATCGTCAA 809  
Db 1716 GGTTCGACGCCATCCGCAATGCGTCAAGCTGGCGGAAAGCTGCGGTGCGCATCGTCAA 1775  
Qy 810 TAACAGTTTTTGAACAAACATCGAGGCGAGGCACTGCCGCCCAATTTCCAAATAGCCAAATC 869  
|||||

Db 1776 TAACAGTTTGGAAACAACATCGAGGGCAGGCACTGCGGACCTTTTCCAAATAGCCAAATTC 1835  
QY 870 GGAGGAGCACTACCGCCAAAGCGTTGCTCGCCTATTTCGGCGGTGATATAACAGACAGAGGG 929  
Db 1836 GGAGGAGCACTACCGCCAAAGCGTTGCTCGACTATTTCGGCGGTGATATAACAGACAGAGGG 1895  
QY 930 TATCCGGCTGATGCACAGAGCGATTACGGCACTTGTCTTACACATCCGTATAAAAA 989  
Db 1896 TATCCGGCTGATGCACAGAGCGATTACGGCAACCTGTCTTACACATCCGTATAAAAA 1955  
QY 990 CATGCTTTTTCATTTTTCGGCAAGCAATGACGCACAAAGCTCAGCCCAACACACTGACCCT 1049  
Db 1956 CATGCTTTTTCATTTTTCGGACAGCAATGACGCACAAAGCTCAGCCCAACACATGACCCT 2015  
QY 1050 ATTGCCATTTTATGAAAAAGATGCTCAAAAAGGCATTATCACAGTCGACGGCGTAGACCG 1109  
Db 2016 ATTGCCATTTTATGAAAAAGCCTCAAAAAGGCATTATCACAGTCGACGGCGTAGACCG 2075  
QY 1110 CAGTGGAGAAAAGTTCA-----A 1127  
Db 2076 CAGTGGAGAAAAGTTCAAAACGGGAAATGTATGGAGAAACCGGTTACGAAACCGCTTGAGTA 2135  
QY 1128 TGGCTCCAACCATTTGGGGAATTACTGCCATGTGTGCTATTCGSCACCCCTATGAAGCAAG 1187  
Db 2136 TGGCTCCAACCATTTGGGGAATTACTGCCATGTGTGCTGTCGGCACCCCTATGAAGCAAG 2195  
QY 1188 CGTCCGTTTTCACCCGTACAAACCCGATTCAAATTTGCCGGAACATCCTTTTCCGCAACCCAT 1247  
Db 2196 CGTCCGTTTTCACCCGTACAAACCCGATTCAAATTTGCCGGAACATCCTTTTCCGCAACCCAT 2255  
QY 1248 CGTAACCGGACCGGGCTCTGCTGTGTCAGAAATACCCGTGATGAGCAACGACAACT 1307  
Db 2256 CGTAACCGGACCGGGCTCTGCTGTGTCAGAAATACCCGTGATGAGCAACGACAACT 2315  
QY 1308 GCCTACCACTGCTGCACACGCTCAGGACATCGGTGCACTGGCGTGACAGCAAGTT 1367  
Db 2316 GCCTACCACTGCTGCACAGCGCTCAGGACATCGGTGCACTGGCGTGACAGCAAGTT 2375  
QY 1368 CGCTGGGACTGCTGGATCGGGTAAGGCCATGAAGCCGCGCTTCCTTCGG 1427  
Db 2376 CGCTGGGACTGCTGGATCGGGTAAGGCCATGAAGCCGCGCTTCCTTCGG 2435  
QY 1428 CGACTTTACCGCGATACGAAAGGTACATCCGATATTGCTACTCTCTTCGTTAACGACAT 1487  
Db 2436 CGACTTTACCGCGATACGAAAGGTACATCCGATATTGCTACTCTCTTCGTTAACGACAT 2495  
QY 1488 TTCAGCACGGCGGCTGTATCAAAAAGCGCGCACCACTGCAACTGCACGGCACAA 1547  
Db 2496 TTCAGCACGGCGGCTGTATCAAAAAGCGCGCACCACTGCAACTGCACGGCACAA 2555  
QY 1548 CACTTACGGGCAAAACCATTTATCGAAGCGGTTTCGCTGGTGTGTGTACGGCAACACAA 1607  
Db 2556 CACTTACGGGCAAAACCATTTATCGAAGCGGTTTCGCTGGTGTGTGTACGGCAACACAA 2615  
QY 1608 ATCGGATATCGCGTCAAAACAAAGGTGCGCTGATTTATAACGGGCGGCATCCGGCGG 1667  
Db 2616 ATCGGATATCGCGTCAAAACAAAGGTGCGCTGATTTATAACGGGCGGCATCCGGCGG 2675  
QY 1668 TAGCCTGAACAGCAGCGGCACTGCTATCTGCGCAGATACCGACCGATCCGGCGCAACGA 1727  
Db 2676 CACCTGAACAGCAGCGGCACTGCTATCTGCGCAGATACCGACCAATCCGGCGCAACGA 2735  
QY 1728 AACCGTGCAATCAAAAGCGGATCGAGCTGGGCGGCAAGGTACGCTGTACACAGTTT 1787  
Db 2736 AACCGTGCAATCAAAAGCGGATCGAGCTGGACGGCAAGGTACGCTGTACACAGTTT 2795  
QY 1788 GGGCAACTGCTGAAAGTGACCGGTACGGGGATGACCGGGCGGAAGGTGTACATGTGCGG 1847  
Db 2796 GGGCAACTGCTGAAAGTGACCGGTACGGGGATGACCGGGCGGAAGGTGTACATGTGCGG 2855  
QY 1848 ACCGGCAAGGGCGAGGCTATCTCAACCGTACCGGACACAGTTTCCCTTCCTGAGTGC 1907  
Db 2856 ACCGGCAAGGGCGAGGCTATCTCAACAGTACCGGACACAGTTTCCCTTCCTGAGTGC 2915

QY 1908 CGCCAAATCGGGCGGATTTATCTTTCTTCAAAACATCGAAACCGACGCTGCTGTGCT 1967  
Db 2916 CGCCAAATCGGGCAGGATTTATCTTTCTTCAAAACATCGAAACCGACGGCGCCTGCT 2975  
QY 1968 GGGTTCCCTCGACAGCTCGAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCTCTATTA 2027  
Db 2976 GGGTTCCCTCGACAGCTCGAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCTCTATTA 3035  
QY 2028 TGTCCGTCGGGCAATCGGCACGGACTGCTTCGGCAGCGGCACATTCGCGGCCCGCGG 2087  
Db 3036 TGTCCGTCGGGCAATCGGCACGGACTGCTTCGGCAGCGGCACATTCGCGGCCCGCGG 3095  
QY 2088 TCTGAAACACCGCGTAGAACAGGGCGGCAGCAATCTCGAAAAACCTGATGTCGAACCTGGA 2147  
Db 3096 TCTGAAACACCGCGTAGAACAGGGCGGCAGCAATCTCGAAAAACCTGATGTCGAACCTGGA 3155  
QY 2148 TGGCTCGAATCATPCCGCAACACCGGAGAGCGGTTGAAACTGCGGCCCGCCGACGACAG 2207  
Db 3156 TGGCTCGAATCATPCCGCAACACCGGAGAGCGGTTGAAACTGCGGCAGCGCCGACGACAG 3215  
QY 2208 TATGCCGGGATCGGCCCTACGGCGCAACTTCCGCGCAGCGGCACGCTACAGCATGC 2267  
Db 3216 TATGCCGGGATCGGCCCTACGGCGCAACTTCCGCGCAGCGGCACGCTACAGCATGC 3275  
QY 2268 GAATGCCCGCAGCGGTGACGCATCTTCAACAGTCTCGCCGCTACCGCTCTATGCCGACAG 2327  
Db 3276 GAATGCCCGCAGCGGTGACGCATCTTCAACAGTCTCGCCGCTACCGCTCTATGCCGACAG 3335  
QY 2328 TACCGCGCCCATGCCGATATGCAAGGACCGCGCTGAAACCGGTATCGACGGGTTGGA 2387  
Db 3336 TACCGCGCCCATGCCGATATGCAAGGACCGCGCTGAAACCGGTATCGACGGGTTGGA 3395  
QY 2388 CCACAGCTTACGGGTGCGCGTCTACGCGCAATCGCGCAACCCACAGACGCTGGAAGTGGGA 2447  
Db 3396 CCACAGCTTACGGGTGCGCGTCTACGCGCAATCGCGCAACCCACAGACGCTGGAAGTGGGA 3455  
QY 2448 ACAGGGCGGTGTTGAAGGCATAATGCGCGCAGTACCCAAACCGCTCGGCATTTGCCGGCAA 2507  
Db 3456 ACAGGGCGGTGTTGAAGGCATAATGCGCGCAGTACCCAAACCGCTCGGCATTTGCCGGCAA 3515  
QY 2508 AACCGGGAAAAATACGACAGACGCGCCACACTGGGCATGGGACACAGCACATGGAGCGA 2567  
Db 3516 AACCGGGAAAAATACGACAGACGCGCCACACTGGGCATGGGACACAGCACATGGAGCGA 3575  
QY 2568 AAACAGTGAATGCAAAAACCGACACATAGTCTGCTTTGCAAGGCATACGGCACGATGC 2627  
Db 3576 AAACAGTGAATGCAAAAACCGACACATAGTCTGCTTTGCAAGGCATACGGCACGATGC 3635  
QY 2628 GGGCGATATCGGCTATCTCAAAGCGCTGTTCTCCTACGGACGCTTACAAAAACAGCATCAG 2687  
Db 3636 GGGCGATATCGGCTATCTCAAAGCGCTGTTCTCCTACGGACGCTTACAAAAACAGCATCAG 3695  
QY 2688 CCGCAGACCGGTGCGGACGAACATCGGAAGGACGGTCAACGGCACGCTGATGACGCT 2747  
Db 3696 CCGCAGACCGGTGCGGACGAACATCGGAAGGACGGTCAACGGCACGCTGATGACGCT 3755  
QY 2748 GGGCGACTGGGCGGTCTCAACGTTCCGTTTGGCGAACGGGAGATTGACGGTCAAGG 2807  
Db 3756 GGGCGACTGGGCGGTCTCAACGTTCCGTTTGGCGAACGGGAGATTGACGGTCAAGG 3815  
QY 2808 CGGTCTCGCTACGACCTGCTCAACAGGATGCAATTCGCCGAAAAAGGCAAGTGTCTTGGG 2867  
Db 3816 CGGTCTCGCTACGACCTGCTCAACAGGATGCAATTCGCCGAAAAAGGCAAGTGTCTTGGG 3875  
QY 2868 CTGGAGGGCAACAGCCTCACTGAAGGACACACTGGTTCGACTCGCGGCTCTGAAGCTGTC 2927  
Db 3876 CTGGAGGGCAACAGCCTCACTGAAGGACACGCTGGTTCGACTCGCGGCTCTGAAGCTGTC 3935  
QY 2928 GCAACCTTTGAGCGATAAAGCCGCTCTGTTTGCAACGGCGGGCGTGAACCGCACCTGAA 2987  
Db 3936 GCAACCTTTGAGCGATAAAGCCGCTCTGTTTGCAACGGCGGGCGTGAACCGCACCTGAA 3995



```
QY 2988 CGGACGCGACTACACGGTAACGGCGCGCTTTTACCGCGCGACTGCGAGCAACCGGCAAGAC 3047
Db 3996 CGGACGCGACTACACGGTAACGGCGCGCTTTTACCGCGCGACTGCGAGCAACCGGCAAGAC 4055
QY 3048 GGGGGACCAATATGCCGCACACCGCCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATT 3107
Db 4056 GGGGGACCAATATGCCGCACACCGCCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATT 4115
QY 3108 CGGCAACGCTTGAACGGCTTGGCAGCTTACAGCTACGCGCGTTCCAAACAGTACGGCAA 3167
Db 4116 CGGCAACGCTTGAACGGCTTGGCAGCTTACAGCTACGCGCGTTCCAAACAGTACGGCAA 4175
QY 3168 CCACAGCGACGAGTGGCGGTAGGCTACCGGTTCTGA 3204
Db 4176 CCACAGCGACGAGTGGCGGTAGGCTACCGGTTCTGA 4212

RESULT 9
AX236457
LOCUS AX236457 4335 bp DNA linear PAT 26-SEP-2001
DEFINITION Sequence 150 from Patent WO0164922.
ACCESSION AX236457
VERSION AX236457.1 GI:15796050
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Arico,M.B., Comanducci,M.C., Galeotti,C.C., Masignani,V.C.,
Guillani,M.M. and Pizzi,M.C.
TITLE Heterologous expression of neisserial proteins
JOURNAL Patent: WO 0164922-A 150 07-SEP-2001;
Chiron spa (It)
FEATURES
Source
1..4335
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="961-983"
BASE COUNT 1217 a 1189 c 1165 g 764 t
ORIGIN
Query Match 93.1%; Score 2983.8; DB 6; Length 4335;
Best Local Similarity 96.9%; Pred. No. 0;
Matches 3070; Conservative 0; Mismatches 57; Indels 42; Gaps 1;
QY 75 TTCTGCTTGTAGCGCGCGCGGCGACTTCTGCGCCGACTTCAATGCGAGCGGCAC 134
Db 1140 TTACGAGTGGGATCGCGGAGGCGGCGACTTCTGCGCCGACTTCAATGCGAGCGGTAC 1199
QY 135 CGGTATCGGACGACGAGCAGCAACACAGCGAAATCAGCAGCAGTATCTTACGCGG 194
Db 1200 CGGTATCGGACGACGAGCAGCAACACAGCGAAATCAGCAGCAGTATCTTACGCGG 1259
QY 195 TATCAAGAACGAATGTGAAAGACAGAGCATGCTCTGTGCGGTCGGGATGACGTTGC 254
Db 1260 TATCAAGAACGAATGTGAAAGACAGAGCATGCTCTGTGCGGTCGGGATGACGTTGC 1319
QY 255 GGTACAGACAGGATGCCAAATCAATGCCCGCCCGGAACTCTGCATACCGGAGACTT 314
Db 1320 GGTACAGACAGGATGCCAAATCAATGCCCGCCCGGAACTCTGCATACCGGAGACTT 1379
QY 315 TACAAACCCAAATGAGCGATACAAAGAAATTTGATCAACCTCAACCTGCAATGAAGCAGG 374
Db 1380 TCCAAACCCAAATGAGCGATACAAAGAAATTTGATCAACCTCAACCTGCAATGAAGCAGG 1439
QY 375 CTATACAGACGCGGGGTAGAGTATGCTGATACAGGCGGAATCCGTCGGCAGCAT 434
Db 1440 CTATACAGACGCGGGGTAGAGTATGCTGATACAGGCGGAATCCGTCGGCAGCAT 1499
QY 435 ATCCCTTCCGGAACCTATGGCAGAAAACACAGCGCTATACGAAATTAACAAAACATA 494
Db 1500 ATCCCTTCCGGAACCTATGGCAGAAAACACAGCGCTATACGAAATTAACAAAACATA 1559
```

```
QY 495 TACGCGGTATATCGGAAGAAAGCGCCTGAAGACGAGCGGTAAAGACATTAAAGCTTC 554
Db 1560 TACGCGGTATATCGGAAGAAAGCGCCTGAAGACGAGCGGTAAAGACATTAAAGCTTC 1619
QY 555 TTTCGACGATGAGCGCGTTATAGAGACTGAAGCAAGCCGACGGATATCCGCCACGTAA 614
Db 1620 TTTCGACGATGAGCGCGTTATAGAGACTGAAGCAAGCCGACGGATATCCGCCACGTAA 1679
QY 615 AGAAATCGGACACATGCGATGTGCTCCCATATATTATGGCGGCGTTCGTCGACGGCAG 674
Db 1680 AGAAATCGGACACATGCGATGTGCTCCCATATATTATGGCGGCGTTCGTCGACGGCAG 1739
QY 675 ACCTGACGCGGTATTGCGCCGATGCGACGCTACACATAAATACATCATGATGAAC 734
Db 1740 ACCTGACGCGGTATTGCGCCGATGCGACGCTACACATAAATACATCATGATGAAC 1799
QY 735 CAAGAAGCAATATGCTGCGACCATCGCAATGATGGTCAAGCTGGGCGAAGCTGG 794
Db 1800 CAAGAAGCAATATGCTGCGACCATCGCAATGATGGTCAAGCTGGGCGAAGCTGG 1859
QY 795 CGTGGCATCGTCAATAACAGTTTGGAAACAACATCGAGGCGAGGCACCTGCCGACCATTT 854
Db 1860 CGTGGCATCGTCAATAACAGTTTGGAAACAACATCGAGGCGAGGCACCTGCCGACCATTT 1919
QY 855 CCAATAGCCAAATTGCGAGGAGCAGTACCGCCAAGCGTTGCTCGCCTATTCCGCGCGTGA 914
Db 1920 CCAATAGCCAAATTGCGAGGAGCAGTACCGCCAAGCGTTGCTCGCCTATTCCGCGCGTGA 1979
QY 915 TAAACAGACGAGGATCCGCTGATGCAACAGACGCGATTTACGGCAACTGTCTCTACCA 974
Db 1980 TAAACAGACGAGGATCCGCTGATGCAACAGACGCGATTTACGGCAACTGTCTCTACCA 2039
QY 975 CATCGTAATAAAACATGCTTTTCATTTTTCGGCAAGCAATGACGCAACAGCTCAGCC 1034
Db 2040 CATCGTAATAAAACATGCTTTTCATTTTTCGGCAAGCAATGACGCAACAGCTCAGCC 2099
QY 1035 CAACACACTGACCTATTGCCATTTTATGAAAAGATGCTCAAAAAGCATTTATCACAGT 1094
Db 2100 CAACACATATGCCCTATTGCCATTTTATGAAAAGATGCTCAAAAAGCATTTATCACAGT 2159
QY 1095 CGCAGCGTAGACCGCAGTGGAGAAAGTTCA----- 1126
Db 2160 CGCAGCGTAGACCGCAGTGGAGAAAGTTCAAAACGGGAAATGTATGGAGAACCGGTAC 2219
QY 1127 -----ATGCTTCCAAACCATTTGCGCAATTTACTGCCATGCTGCTCTATCGGC 1172
Db 2220 AGAACCGTTGAGTATGGCTCCAACCATTTGCGCAATTTACTGCCATGCTGCTCTCGGC 2279
QY 1173 ACCCTATGAAGCAAGCGTCCGTTTCAACCGGTACAAACCGGATTCAAATTCGCGGAACATC 1232
Db 2280 ACCCTATGAAGCAAGCGTCCGTTTCAACCGGTACAAACCGGATTCAAATTCGCGGAACATC 2339
QY 1233 CTTTTCGACCCATGCTAACCGGACGCGGCTCTGCTGCTGCGCAAAATACCCGTGGAT 1292
Db 2340 CTTTTCGACCCATGCTAACCGGACGCGGCTCTGCTGCTGCGCAAAATACCCGTGGAT 2399
QY 1293 GAGCAACGACAACTGCGTACCACGCTGCTGACACGGCTCAGGACATCGGTGACGTCGG 1352
Db 2400 GAGCAACGACAACTGCGTACCACGCTGCTGACACGGCTCAGGACATCGGTGACGTCGG 2459
QY 1353 CGTGGACAAAGTTGCGGTGGGACTGCTGGATGCGGGTAAAGGCGCATGAACGACCCGC 1412
Db 2460 CGTGGACAAAGTTGCGGTGGGACTGCTGGATGCGGGTAAAGGCGCATGAACGACCCGC 2519
QY 1413 GTCCCTTCGTTGCGGCACTTTACCGCCGATACGAAAGGTACATCCGATATTCCTTACTC 1472
Db 2520 GTCCCTTCGTTGCGGCACTTTACCGCCGATACGAAAGGTACATCCGATATTCCTTACTC 2579
QY 1473 CTTTCCGTAAACGACATTTACGGCACGGCGGCTGATCAAAAAGGCGGAGCACTGCA 1532
Db 2580 CTTTCCGTAAACGACATTTACGGCACGGCGGCTGATCAAAAAGGCGGAGCACTGCA 2639
QY 1533 ACTGCACGCAACACACCTATACGGGCAAAACCATTTATCGAAGCGGTTTCGCTGTGT 1592
```



[illegible]

D	b	3720	CATACGGCAGCATGCGGGCGATATCGGCTACTCTCAAAGGCCTTGTCTCTCAACGACGCCTA	3777
Q	y	2673	CAAAACAGCATACGCGCAGCACCGTTCGGAGCAACAATCGGAAGCGACGCGTCAACGG	2732
D	b	3780	CAAAAACAGCATACGCGCAGCACCGTTCGGAGCAACAATCGGAAGCGACGCGTCAACGG	3839
Q	y	2733	CACGCTGATCAGTGCGGGCGACTGGCGGGTGTCAACGTTCCGTTTGGCCGAACGGGAGA	2792
D	b	3840	CACGCTGATCAGTGCGGGCGACTGGCGGGTGTCAACGTTCCGTTTGGCCGAACGGGAGA	3899
Q	y	2793	TTTGACGGTCCAAGGGCGTCTGGCTACGACCTGCTCAAACAGGATGATTTCGCCGAAAA	2852
D	b	3900	TTTGACGGTCCAAGGGCGTCTGGCTACGACCTGCTCAAACAGGATGATTTCGCCGAAAA	3959
Q	y	2853	AGGCAGTGCTTTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACGTGTCGACATCGC	2912
D	b	3960	AGGCAGTGCTTTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACGTGTCGACATCGC	4019
Q	y	2913	GGGTCTGAAGCTGTGCGCAACCCCTTGAGCGATAAAGCCGCTCTGTTTGAACGGGGCGGT	2972
D	b	4020	GGGTCTGAAGCTGTGCGCAACCCCTTGAGCGATAAAGCCGCTCTGTTTGAACGGGGCGGT	4079
Q	y	2973	GGAACGGCAGCTGAACGGAGCGGACTACACGGTAAGCGGGCGGCTTTACCGGGCGAGTGC	3032
D	b	4080	GGAACGGCAGCTGAACGGAGCGGACTACACGGTAAGCGGGCGGCTTTACCGGGCGAGTGC	4139
Q	y	3033	AGCAACCGCAAGACGGGGCGACGAATAATGCCGCACACC CGCCTGGTTGCCGGTCTCTGGG	3092
D	b	4140	AGCAACCGCAAGACGGGGCGACGAATAATGCCGCACACC CGCCTGGTTGCCGGCCTTGGG	4199
Q	y	3093	CGCGGATGTCGAATTGCGCAACGGCTGGAACGGCTTGGCAGCTTACAGCTACGCCGGTTC	3152
D	b	4200	CGCGGATGTCGAATTGCGCAACGGCTGGAACGGCTTGGCAGCTTACAGCTACGCCGGTTC	4259
Q	y	3153	CAAAACATACGGCAACACACAGCGGACGAGTGGCGGTAGGCTACCGGTTTC	3201
D	b	4260	CAAAACATACGGCAACACACAGCGGACGAGTGGCGGTAGGCTACCGGTTTC	4308
RESULT 10				
AX239743				
LOCUS AX239743 4335 bp DNA linear PAT 26-SEP-				
DEFINITION Sequence 45 from Patent WO0164920.				
ACCESSION AX239743				
VERSION AX239743.1 GI:15797382				
KEYWORDS synthetic construct.				
SOURCE synthetic construct.				
ORGANISM artificial sequences.				
REFERENCE 1 (bases 1 to 4335)				
AUTHORS Arico,M.B., Comanducci,M., Galeotti,C., Masignani,V., Giuliani,L. and Pizza,M.				
TITLE Hybrid expression of neisserial proteins				
JOURNAL Patent: WO 0164920-A 45 07-SEP-2001;				
Chiron Spa (IT)				
FEATURES				
source Location/Qualifiers				
1..4335				
/organism="synthetic construct"				
/db_xref="taxon:32630"				
/note="961-983"				
BASE COUNT 1217 a 1189 c 1165 g 764 t				
ORIGIN				
Query Match 93.1%; Score 2983.8; DB 6; Length 4335;				
Best Local Similarity 96.9%; Pred. No. 0;				
Matches 3070; Conservative 0; Mismatches 57; Indels 42; Gaps				
Q	y	75	TTCTGCTGCTTAGCGGGCGGGCGGCACTTCTGGCCCCGACTTCATGACGGCGCAC	134
D	b	1140	TTACAGTGGGATCCGGGGAGGGCGCACTTCTGGCCCCGACTTCATGACGGCGTAC	1199
Q	y	135	CGGTATCGGCAGCAACAGCAGACGAACAACCGGAAATCAGCAGCATGATCTTACGCCGG	194

Db 1200 CGGTATCGGACGCAACACAGAGCAACACAGCGAAATACAGCAGCAGTATCTTACGCCGG 1259  
Qy 195 TATCAAGAACGAAATGTGCAAGACAGAGCATGCTCTGTGCGGTTCGGGTGACGTTGC 254  
Db 1260 TATCAAGAACGAAATGTGCAAGACAGAGCATGCTCTGTGCGGTTCGGGTGACGTTGC 1319  
Qy 255 GGTTCAGACAGGGATGCCAAATCAATGCCCCCCCCCGGAAATCTGCAATACCGGAGACTT 314  
Db 1320 GGTTCAGACAGGGATGCCAAATCAATGCCCCCCCCCGGAAATCTGCAATACCGGAGACTT 1379  
Qy 315 TACAAACCAATGACGATACAGAAATTTGATCAACCTCAACCTCAACCTGCAATTAAGCAGG 374  
Db 1380 TCAAAACCAATGACGATACAGAAATTTGATCAACCTCAACCTCAACCTGCAATTAAGCAGG 1439  
Qy 375 CTATACAGACCGGGTAGAGTAGGTATCGTCATACAGCGAATCCGTCGGCAGCAT 434  
Db 1440 CTATACAGACCGGGTAGAGTAGGTATCGTCACACAGCGGAATCCGTCGGCAGCAT 1499  
Qy 435 ATCCCTTCCCGAATCTATGGCAGAAAAAGACACGGCTATACGAAAAATACAAAACTA 494  
Db 1500 ATCCCTTCCCGAATCTATGGCAGAAAAAGACACGGCTATACGAAAAATACAAAACTA 1559  
Qy 495 TACGGCGTATATGCGGAAGGAGCGCTGAAGACGGAGCGGTAAAGACATTAAGCTTC 554  
Db 1560 TACGGCGTATATGCGGAAGGAGCGCTGAAGACGGAGCGGTAAAGACATTAAGCTTC 1619  
Qy 555 TTTCGACGATGAGCGCTTATAGACCTGAAGCAAGCGCAGCGATATCCGCGACGTAA 614  
Db 1620 TTTCGACGATGAGCGCTTATAGACCTGAAGCAAGCGCAGCGATATCCGCGACGTAA 1679  
Qy 615 AGAAATCGGACACATCGATGTGCTCCCATATATTATGGCGGCGTCCGTCGAGCGGAG 674  
Db 1680 AGAAATCGGACACATCGATTTGGTCTCCATATATTATGGCGGCGTCCGTCGAGCGGAG 1739  
Qy 675 ACCTCGCGCGTATGCGCGCGATGCGACGCTACACATATGAATACGATGGAAC 734  
Db 1740 ACCTCGAGCGGTATGCGCGCGATGCGACGCTACACATATGAATACGATGGAAC 1799  
Qy 735 CAAGAACGAAATATGCTGACGCCATCGCAATGCATGGGTCAAGCTGGGGGAACGTGG 794  
Db 1800 CAAGAACGAAATATGCTGCGACCATCGGCAATGCATGGGTCAAGCTGGGGGAACGTGG 1859  
Qy 795 CGTGCATCGTCAATACAGTTTGGAAACACATCGAGGGCAGCATGCGGACCATTT 854  
Db 1860 CGTGCATCGTCAATACAGTTTGGAAACACATCGAGGGCAGCATGCGGACCATTT 1919  
Qy 855 CCAATAGCCAAATCGAGGAGCAGTACCGCAACGCTGCTCGCTATTCGGCGGTGA 914  
Db 1920 CCAATAGCCAAATCGAGGAGCAGTACCGCAACGCTGCTCGCTATTCGGCGGTGA 1979  
Qy 915 TAAACAGACGAGGTATCCGCTGATGCAACAGAGCGATTAAGGCAACTTGTCTACCA 974  
Db 1980 TAAACAGACGAGGTATCCGCTGATGCAACAGAGCGATTAAGGCAACTTGTCTACCA 2039  
Qy 975 CATCGTAAATAACATGCTTTTCATTTTTCGGCAAGCAATGACGCAACAGCTCAGCC 1034  
Db 2040 CATCGTAAATAACATGCTTTTCATTTTTCGACAGCAATGACGCAACAGCTCAGCC 2099  
Qy 1035 CAACACATGACCTTATGCGCAATTTATGAAAGACATGCTCAAAAGGCATTTATCACAGT 1094  
Db 2100 CAACACATGACCTTATGCGCAATTTATGAAAGACATGCTCAAAAGGCATTTATCACAGT 2159  
Qy 1095 CGCAGCGTAGACCGCAGTGGAGAAAGTTCA - ----- 1126  
Db 2160 CGCAGCGTAGACCGCAGTGGAGAAAGTTCAAAACGGGAAATGTATGAGAACCGGTTAC 2219  
Qy 1127 -----ATGGCTCAACCATTTGGGAATTAATGCGCATGTTGGCTATCGGC 1172  
Db 2220 AGAACCGCTTGAATGCTCAACCATTTGGGAATTAATGCGCATGTTGGCTATCGGC 2279  
Qy 1173 ACCCTATAGCAACGCTCGCTTACCCGCTACAAACCGGATTTCAAAATGCGGGAACATC 1232  
Db 2280 ACCCTATAGCAACGCTCGCTTACCCGCTACAAACCGGATTTCAAAATGCGGGAACATC 2339

Qy 1233 CTTTTCGCAACCATCGTAACCGCACGCGGCTCTGCTGCTGCAAAATACCCGTGGAT 1292  
Db 2340 CTTTTCGCAACCATCGTAACCGCACGCGGCTCTGCTGCTGCAAAATACCCGTGGAT 2399  
Qy 1293 GAGCAACGACAACTCGCTACACGCTGCTGACAAAGCGCTCAGGACATCGGTGCACTCGG 1352  
Db 2400 GAGCAACGACAACTCGCTACACGCTGCTGACAAAGCGCTCAGGACATCGGTGCACTCGG 2459  
Qy 1353 CGTGACAGCAAGTTCGGCTGGGACTGCTGATCGCGGTAAAGCCATTAAGCGACCGC 1412  
Db 2460 CGTGACAGCAAGTTCGGCTGGGACTGCTGATCGCGGTAAAGCCATTAAGCGACCGC 2519  
Qy 1413 GTCCTTTCGGCTGGGACTTACCGCGCATACGAAAGGTACATCCGATATTGCTACTC 1472  
Db 2520 GTCCTTTCGGCTGGGACTTACCGCGCATACGAAAGGTACATCCGATATTGCTACTC 2579  
Qy 1473 CTTCCGTAAACGACATTTACGACGCGGCGCTGATCAAAAAGCGCGACGCAACTGCA 1532  
Db 2580 CTTCCGTAAACGACATTTACGACGCGGCGCTGATCAAAAAGCGCGACGCAACTGCA 2639  
Qy 1533 ACTGCACGGCAACACACTATACGGGCAAAACCATTTATCAAGCGGTTGCTGGTGT 1592  
Db 2640 ACTGCACGGCAACACACTATACGGGCAAAACCATTTATCAAGCGGTTGCTGGTGT 2699  
Qy 1593 GTACGGCAACCAAAATCGGATATGCGCTCGAAACCAAGGTGCTGATTATTAACGG 1652  
Db 2700 GTACGGCAACCAAAATCGGATATGCGCTCGAAACCAAGGTGCTGATTATTAACGG 2759  
Qy 1653 GCGGCAATCGGCGGTAGCTGAAACAGCAGCGCATTTCTATCTGGCAGATACGACCG 1712  
Db 2760 GCGGCAATCGGCGCGCTGAAACAGCAGCGCATTTCTATCTGGCAGATACGACCA 2819  
Qy 1713 ATCGGCGCAACCAACCGTGCACATCAAAAGCGCATCTGCAGCTGGCGGCGGAGTAC 1772  
Db 2820 ATCGGCGCAACCAACCGTGCACATCAAAAGCGCATCTGCAGCTGGCGGCGGAGTAC 2879  
Qy 1773 GCTGTACACAGCTTTGGGCAAACTGCTGAAAGTGACGCTACGGCGATTAATCGCGGCA 1832  
Db 2880 GCTGTACACAGCTTTGGGCAAACTGCTGAAAGTGACGCTACGGCGATTAATCGCGGCA 2939  
Qy 1833 GCTGTACATGTCGGCACCGGCAAAAGGGCAGGCTATCTCAACCGTACCGGCAACGCT 1892  
Db 2940 GCTGTACATGTCGGCACCGGCAAAAGGGCAGGCTATCTCAACAGTACCGGCAACGCT 2999  
Qy 1893 TCCCTTCTGTAGTCCGCCAAAATCGGCGGATTTCTTTTTCACAAACATCGAAAC 1952  
Db 3000 TCCCTTCTGTAGTCCGCCAAAATCGGCGAGATTTCTTTTTCACAAACATCGAAAC 3059  
Qy 1953 CGACGGTGTCTGCTGGCTTCCCTCGACAGCTCGAAACACAGCGGCGAGTGAAGCGA 2012  
Db 3060 CGACGGCGCTGCTGGCTTCCCTCGACAGCTCGAAACACAGCGGCGAGTGAAGCGA 3119  
Qy 2013 CACGCTGTCTTATGTCGTCGGGCAATGCGGCAGGACTGCTTCGGCAGCGGCA 2072  
Db 3120 CACGCTGTCTTATGTCGTCGGGCAATGCGGCAGGACTGCTTCGGCAGCGGCA 3179  
Qy 2073 TTCGCGCGCGCTGTGAAACACGCGTAGAACAGGGCGGCAATCTGGAACCT 2132  
Db 3180 TTCGCGCGCGCGCTGTGAAACACGCGTAGAACAGGGCGGCAATCTGGAACCT 3239  
Qy 2133 GATGCTCGAACTGGATGCTCCGAATCATCGCAACACCCGAGAGGTTGAACCTGCGC 2192  
Db 3240 GATGCTCGAACTGGATGCTCCGAATCATCGCAACACCCGAGAGGTTGAACCTGCGC 3299  
Qy 2193 CGCGACCGCAGATATGCGGCGATCCGCCCTACGGCGCACTTTCGGCGAGCGC 2252  
Db 3300 AGCGACCGCAGATATGCGGCGATCCGCCCTACGGCGCACTTTCGGCGAGCGC 3359  
Qy 2253 AGCGTACAGCATGGAATGCGCGAGCGTGTACGCACTTTCAACAGTCTCGCGCTAC 2312  
Db 3360 AGCGTACAGCATGGAATGCGCGAGCGTGTACGCACTTTCAACAGTCTCGCGCTAC 3419

QY 2313 CGTCTATGCGACGATACCGCGCCCATCCGATATGCAAGGACGCGCGCTGAAAGCCGT 2372  
Db 3420 CGTCTATGCGACGATACCGCGCCCATCCGATATGCAAGGACGCGCGCTGAAAGCCGT 3479  
QY 2373 ATCGGACGGTTGGACCAACAGCTACGGGTCTGCGCTATCGCCCAACCCACAGGA 2432  
Db 3480 ATCGGACGGTTGGACCAACAGCGGAGCGGTCTGCGCTATCGCCCAACCCACAGGA 3539  
QY 2433 CGGTGAACGCTGGGAACAGCGGCTGTTGAAGGCAAAATGCGCGGAGTACCCAAACCGT 2492  
Db 3540 CGGTGAACGCTGGGAACAGCGGCTGTTGAAGGCAAAATGCGCGGAGTACCCAAACCGT 3599  
QY 2493 CGGCATTCGCGGAAACCGGCGAAATACGACAGCGCGCCACACTGGGATGGGACA 2552  
Db 3600 CGGCATTCGCGGAAACCGGCGAAATACGACAGCGCGCCACACTGGGATGGGAGC 3659  
QY 2553 CAGCACATGGAGGAAACAGTGCATATCAAAATCCAAACCGACAGTATAGTCTGTTGCAGG 2612  
Db 3660 CAGCACATGGAGGAAACAGTGCATATCAAAATCCAAACCGACAGTATAGTCTGTTGCAGG 3719  
QY 2613 CATACGGCACGATGCGGGCGATATCGGCTATCTCAAAGCCCTGTTCTCTACGGACGCTA 2672  
Db 3720 CATACGGCACGATGCGGGCGATATCGGCTATCTCAAAGCCCTGTTCTCTACGGACGCTA 3779  
QY 2673 CAAAAACAGCATCAGCGCGACGACCGGTGCGGACGACATGGGAGGAGCGTCAACGG 2732  
Db 3780 CAAAAACAGCATCAGCGCGACGACCGGTGCGGACGACATGGGAGGAGCGTCAACGG 3839  
QY 2733 CAGCTGATGCACTGGGCGCACTGGGCGGTGTCACGTTCCGTTTGGCGCAACGGGAGA 2792  
Db 3840 CAGCTGATGCACTGGGCGCACTGGGCGGTGTCACGTTCCGTTTGGCGCAACGGGAGA 3899  
QY 2793 TTTGACGGTCGAAGGGGTCTGCGCTACGACCTGCTCAAACAGGATGCAATCGCCGAAA 2852  
Db 3900 TTTGACGGTCGAAGGGGTCTGCGCTACGACCTGCTCAAACAGGATGCAATCGCCGAAA 3959  
QY 2853 AGGCACTGCTTTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTGGACTCGC 2912  
Db 3960 AGGCACTGCTTTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTGGACTCGC 4019  
QY 2913 GGGTCTGAAGCTGTCGCAACCTTGAGCGATTAAGCCGTCCTGTTTGAACGCGGCGCT 2972  
Db 4020 GGGTCTGAAGCTGTCGCAACCTTGAGCGATTAAGCCGTCCTGTTTGAACGCGGCGCT 4079  
QY 2973 GGAACGCACTGAACGCGGCACTACACGCTTAACGCGGCTTTACCGCGCGACTGC 3032  
Db 4080 GGAACGCACTGAACGCGGCACTACACGCTTAACGCGGCTTTACCGCGCGACTGC 4139  
QY 3033 AGCAACGGCAAGCGGGGCAAGCAATATGCGGCACACCCGCTGGTTGCCGGTCTGGG 3092  
Db 4140 AGCAACGGCAAGCGGGGCAAGCAATATGCGGCACACCCGCTGGTTGCCGGTCTGGG 4199  
QY 3093 CCGGATGTCGAATTCGGCAACGGCTGGAACGCTTGGACGCTTGGACGCTACAGCTACCGGGTTC 3152  
Db 4200 CCGGATGTCGAATTCGGCAACGGCTGGAACGCTTGGACGCTTGGACGCTACAGCTACCGGGTTC 4259  
QY 3153 CAAACAGTACGGCAACCAACGCGGACGAGTTCGGCGGTAGGCTACCGGTTCC 3201  
Db 4260 CAAACAGTACGGCAACCAACGCGGACGAGTTCGGCGGTAGGCTACCGGTTCC 4308

RESULT 11  
AX236419  
LOCUS AX236419 3939 bp DNA linear PAT 26-SEP-2001  
DEFINITION Sequence 112 from Patent WO0164922.  
ACCESSION AX236419  
VERSION AX236419.1 GI:15796032  
KEYWORDS  
SOURCE synthetic construct.  
ORGANISM synthetic construct.  
REFERENCE 1 (bases 1 to 3939)  
AUTHORS Arico,M.B., Comanducci,M.C., Galeotti,C., Masignani,I.V.C.,

Guiliani,M. and Pizsa,M.C.  
Heterologous expression of neisserial proteins  
Patent: WO 0164922-A 112 07-SEP-2001;  
Chiron Spa (It)

FEATURES  
Location/Qualifiers  
1..3939  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="deltaG741-983"  
BASE COUNT 1046 a 1103 c 1106 g 684 t  
ORIGIN

Query Match 93.1%; Score 2983.2; DB 6; Length 3939;  
Best Local Similarity 97.1%; Pred. No. 0;  
Matches 3064; Conservative 0; Mismatches 48; Indels 42; Gaps 1;

QY 90 CGCGCGCGCGGACCTTCTGCGCCCGACTTCAATGAGGCGGACCGGTATCGGACGAA 149  
Db 759 CGCGCGGAGCGGACCTTCTGCGCCCGACTTCAATGAGGCGGTAACGGGTATCGGACGAA 818  
QY 150 CAGCAGAGCAACAACAGCGGAAATCAGCAGCAGTATCTTACGCGGTATCAAGAACGAAAT 209  
Db 819 CAGCAGAGCAACAACAGCGGAAATCAGCAGCAGTATCTTACGCGGTATCAAGAACGAAAT 878  
QY 210 GTCAAAGACAGAGCATGCTGTGCGCGTCCGGATGACGTTGCGGTTACAGACAGGGA 269  
Db 879 GTCAAAGACAGAGCATGCTGTGCGCGTCCGGATGACGTTGCGGTTACAGACAGGGA 938  
QY 270 TGC AAAATCAATGCCCCCCCCGAAATCTGCATACCGGAGACTTTTACAAACCCAAATGA 329  
Db 939 TGC AAAATCAATGCCCCCCCCGAAATCTGCATACCGGAGACTTTTCCAAACCCAAATGA 998  
QY 330 CCATACAGAAATTTGATCAACCTCAACCTGCAATTAAGAGAGGCTATACAGGACGCG 389  
Db 999 CGCATACAAGATTTGATCAACCTCAACCTGCAATTAAGAGAGGCTATACAGGACGCG 1058  
QY 390 GGTAGAGGTAGGTATCGTGCATACAGCGGAATCCGTCGGCAGCATATCCTTTCCCGAACT 449  
Db 1059 GGTAGAGGTAGGTATCGTGCATACAGCGGAATCCGTCGGCAGCATATCCTTTCCCGAACT 1118  
QY 450 GTATGCGCAAAAAGACACGGCTTAAACGAAATTAACAAAATTAACGCGGTATATGCG 509  
Db 1119 GTATGCGCAAAAAGACACGGCTTAAACGAAATTAACAAAATTAACGCGGTATATGCG 1178  
QY 510 GAAGGAAGCGCTTGAAGACGGGCGTAAAGACATTAAGACTTCTTTGACGACGAGGC 569  
Db 1179 GAAGGAAGCGCTTGAAGACGGGCGTAAAGACATTAAGACTTCTTTGACGACGAGGC 1238  
QY 570 CGTTATAGAGCTGAAGCAAGCCGACGGATATCCGCCACGTAAAGAAATCGGACACAT 629  
Db 1239 CGTTATAGAGCTGAAGCAAGCCGACGGATATCCGCCACGTAAAGAAATCGGACACAT 1298  
QY 630 CGATGTGGTCTCCCATATTTATTTGGCGGGGCTTCCGTGGAGCGGACACCTGCGAGCGGTAT 689  
Db 1299 CGATTTGGTCTCCCATATTTATTTGGCGGGGCTTCCGTGGAGCGGACACCTGCGAGCGGTAT 1358  
QY 690 TCGCGCCGATGCGACGCTACACATATGAATGACGATGATGAACCAAGCAAGCAATAAT 749  
Db 1359 TCGCGCCGATGCGACGCTACACATATGAATGACGATGATGAACCAAGCAAGCAATAAT 1418  
QY 750 GTCTGCAGCCATCCGCAATGTCATGGGTCAAGCTGGCGGAACGTTGGCGGTGCGCATCTCAA 809  
Db 1419 GGTTCAGCCATCCGCAATGTCATGGGTCAAGCTGGCGGAACGTTGGCGGTGCGCATCTCAA 1478  
QY 810 TAACAGTTTGGAAACAACATCAGGCGGACGACTGCCGACCATTTTCCAAATAGCCAAATTC 869  
Db 1479 TAACAGTTTGGAAACAACATCAGGCGGACGACTGCCGACCTTTTCCAAATAGCCAAATTC 1538  
QY 870 GGAGGAGCAGTACCGCCCAAGCGTTGCTCCCTATTTCCGCGGCTGATAAACAACAGCAGGG 929  
Db 1539 GGAGGAGCAGTACCGCCCAAGCGTTGCTCCCTATTTCCGCGGCTGATAAACAACAGCAGGG 1598  
QY 930 TATCCGCTGATGCAACAGAGCGGATTTACGGCAACTTGTCTACCATCCGTAATAAAAAA 989

Db 1599 TATCCGCTGATGCAACAGAGCGATTACGGCAACCTGCTCTACCACATCCGTAATAAAAA 1658  
Qy 990 CATGCTTTTCATTTTTCGGCAAGCAATGACGCACAAGCTCAGCCCAACACACTGACCC 1049  
Db 1659 CATGCTTTTCATCTTTTCGACAGGCAATGACGCACAAGCTCAGCCCAACACATATGCCCT 1718  
Qy 1050 APTGCCATTTTATGAAAAGATGCTCAAAAAGGCAATATACAGTCGCGAGGCGTAGACGG 1109  
Db 1719 ATTGCCATTTTATGAAAAGAGGCTCAAAAAGGCATATACAGTCGCGAGGCGTAGACGG 1778  
Qy 1110 CAGTGGAGAAAGTTCA-----A 1127  
Db 1779 CAGTGGAGAAAAGTTCAAAACGGGAATGATGGAGAAACCGGGTACAGAACCCGCTTGAAGTA 1838  
Qy 1128 TGGCTCCAAACCAATTCGGAATTTACTGCCATGTTGGTGCCTATCGGCACCCCTATGAAGCAAG 1187  
Db 1839 TGGCTCCAAACCAATTCGGAATTTACTGCCATGTTGGTGCCTGTCGGCACCCCTATGAAGCAAG 1898  
Qy 1188 CGTCCGTTTTCACCCGTACAAACCCGATTCAAAATTCGCCGAACATCTTTTCGCAACCCAT 1247  
Db 1899 CGTCCGTTTTCACCCGTACAAACCCGATTCAAAATTCGCCGAACATCTTTTCGCAACCCAT 1958  
Qy 1248 CGTAACCGCGCAGCGGCTCTGCTGCTGAGAAATACCCGTGGATGAGCAAGCAAACT 1307  
Db 1959 CGTAACCGCGCAGCGGCGCTCTGCTGCTGAGAAATACCCGTGGATGAGCAAGCAAACT 2018  
Qy 1308 GGGTACCACGCTGCTGACAAAGGCTCAGGACATCCGTTGCAGTCGGCGTGAGCAGCAAGTT 1367  
Db 2019 GGGTACCACGCTGCTGACAGCGGCTCAGGACATCCGTCGATCGTGGCGTGAGCAGCAAGTT 2078  
Qy 1368 CGGCTGGGACTGCTGGATGCGGGTAAGGCCATGAACGAGACCCGCTCTTTTCGTTCCG 1427  
Db 2079 CGGCTGGGACTGCTGGATGCGGGTAAGGCCATGAACGAGACCCGCTCTTTTCGTTCCG 2138  
Qy 1428 CGACTTTACCGCGCATACGAAGGTACATCCGATATTCGCTACTCTCTTCGTTAAGCAGAT 1487  
Db 2139 CGACTTTACCGCGCATACGAAGGTACATCCGATATTCGCTACTCTCTTCGTTAAGCAGAT 2198  
Qy 1488 TTCAGGCAAGCGCGCTGATCAAAAAAGGCGGAGCAACTGCAACTGCAGCGGCAACAA 1547  
Db 2199 TTCAGGCAAGCGCGCTGATCAAAAAAGGCGGAGCAACTGCAACTGCAGCGGCAACAA 2258  
Qy 1548 CACCTATACGGCAAAACCATATACGAAGGCGGTTTCGTTGTTGTATCGGCAACAA 1607  
Db 2259 CACCTATACGGCAAAACCATATACGAAGGCGGTTTCGTTGTTGTATCGGCAACAA 2318  
Qy 1608 ATCGGATATGCGCGTCGAACCAAGGTGCGCTGATTTATACGGGGCGGATCCGGCGG 1667  
Db 2319 ATCGGATATGCGCGTCGAACCAAGGTGCGCTGATTTATACGGGGCGGATCCGGCGG 2378  
Qy 1668 TAGCCTGAACAGCGAGCGATTGCTATCTGGCAGATACCGACCATCCGGCGCAACGA 1727  
Db 2379 CAGCCTGAACAGCGAGCGATTGCTATCTGGCAGATACCGACCAATCCGGCGCAACGA 2438  
Qy 1728 AACCTGCAATCAAGGGGATCTCAGCTGGGCGGCAAGGTACGCTGTACACAGTTT 1787  
Db 2439 AACCTGCAATCAAGGGGATCTCAGCTGGGCGGCAAGGTACGCTGTACACAGTTT 2498  
Qy 1788 GGGCAAACTGCTGAAAGTGGAGGTACGGGATGACCGCGCGCAAGCTGTACATCTCGG 1847  
Db 2499 GGGCAAACTGCTGAAAGTGGAGGTACGGGATTTATCGCGGCAAGCTGTACATCTCGG 2558  
Qy 1848 AGCGGCAAGGGGAGGCTATCTCAACCGTACCGGCAACGTTTCCCTTCCCTGAGTGC 1907  
Db 2559 AGCGGCAAGGGGAGGCTATCTCAACAGTACCGGAGCAGCTGTTCCCTTCCCTGAGTGC 2618  
Qy 1908 CGCCAAAATCGGGCGGATATCTTTCTTCAAAAATCGAAAACCGAGGTGCTGCT 1967  
Db 2619 CGCCAAAATCGGGCAGGATATCTTTCTTCAAAAATCGAAAACCGAGGTGCTGCT 2678  
Qy 1968 GGGTTCCTTCGACAGCTTCGAAAAACAGCGGGCAGTGAAGCGCACAGCTGTCTTATTA 2027

Db 2679 GGGTTCCTTCGACAGCGCTGAAAAAACAGCGGGCAGTGAAGGCGACAGCTGTCTCTATTA 2738  
Qy 2028 TGTCCGTCGCGCAATCGGCACAGGACTGCTTCGCGACGCGCACATTCGCGCCCGCGG 2087  
Db 2739 TGTCCGTCGCGCAATCGGCACAGGACTGCTTCGCGACGCGCACATTCGCGCCCGCGG 2798  
Qy 2088 TCTGAAACACACCGCTAGAACAGGCGGCGAGCAATCTGAAAAACCTGATGGTCTGAACTGGA 2147  
Db 2799 TCTGAAACACACCGCTAGAACAGGCGGCGAGCAATCTGAAAAACCTGATGGTCTGAACTGGA 2858  
Qy 2148 TGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTTGGGGCGCGGACCGCACAGA 2207  
Db 2859 TGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTTGGGGCGCGGACCGCACAGA 2918  
Qy 2208 TATGCCGGGCTACCCCTCTAGGGCGCACTTTTCGCGCAGCGGCGAGCTACACATGC 2267  
Db 2919 TATGCCGGGCTACCCCTCTAGGGCGCACTTTTCGCGCAGCGGCGAGCTACACATGC 2978  
Qy 2268 GAATGCCCGCGAGCGTGTACGCATCTTCAACAGTCTCCCGCTACCGTCTATGCCGACAG 2327  
Db 2979 GAATGCCCGCGAGCGTGTACGCATCTTCAACAGTCTCCCGCTACCGTCTATGCCGACAG 3088  
Qy 2328 TACGCCCGCCCATGCCGATATGCAGGAGCGCGGCTGAAAGCCGTATCGGACGGTTTGA 2387  
Db 3039 TACGCCCGCCCATGCCGATATGCAGGAGCGCGGCTGAAAGCCGTATCGGACGGTTTGA 3098  
Qy 2388 CCACAACGCTACGGGTCTGCGGCTCATCGCGCAAAACCCAAACAGGACGGTGAACGTGGGA 2447  
Db 3099 CCACAACGCGAGGCTCTGCGGCTCATCGCGCAAAACCCAAACAGGACGGTGAACGTGGGA 3158  
Qy 2448 ACAGGGCGGTGTTGAAGCAAAATGCGCGGAGTACCCAAACCCGCTGCGCATTTGCCGCGAA 2507  
Db 3159 ACAGGGCGGTGTTGAAGCAAAATGCGCGGAGTACCCAAACCCGCTGCGCATTTGCCGCGAA 3218  
Qy 2508 AACCGGCAAAATACGACAGCAGCGCCACACTGGGCGATGGGACACACATGAGGCGA 2567  
Db 3219 AACCGGCAAAATACGACAGCAGCGCCACACTGGGCGATGGGACACACATGAGGCGA 3278  
Qy 2568 AACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGCGACGATGC 2627  
Db 3279 AACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGCGACGATGC 3338  
Qy 2628 GGGCGATATCGGCTATCTCAAAGGCTCTTCTCTACGGAGCTACAAAAACAGCATCAG 2687  
Db 3339 GGGCGATATCGGCTATCTCAAAGGCTCTTCTCTACGGAGCTACAAAAACAGCATCAG 3398  
Qy 2688 CCGAGCAGCCGTCGGGAGACATGCGGAAGGCGCTCAACGCGACGCTGATCGACT 2747  
Db 3399 CCGAGCAGCCGTCGGGAGCAACATGCGGAAGGCGCTCAACGCGACGCTGATCGACT 3458  
Qy 2748 GGGCGCATGGGCGGTGTCAACGTTCCGTTTCCGCAACGGGAGATTTCAGCGTCAAGG 2807  
Db 3459 GGGCGCATGGGCGGTGTCAACGTTCCGTTTCCGCAACGGGAGATTTCAGCGTCAAGG 3518  
Qy 2808 CGGTCTGCGCTACGACCTGTCAAACAGGATGCAATTCGCCGAAAAAGGCACTGCTTTGGG 2867  
Db 3519 CGGTCTGCGCTACGACCTGTCAAACAGGATGCAATTCGCCGAAAAAGGCACTGCTTTGGG 3578  
Qy 2868 CTGGAGCGCAACAGCTCACTGAAGGCACACTGGTGGGACTCGGGGCTCTGAAGCTGTC 2927  
Db 3579 CTGGAGCGCAACAGCTCACTGAAGGCACACTGGTGGGACTCGGGGCTCTGAAGCTGTC 3638  
Qy 2928 GCAACCTTTGAGCGATAAAGCGCTCTCTTTTGAACGGCGGCGTGAACGCGACTGAA 2987  
Db 3639 GCAACCTTTGAGCGATAAAGCGCTCTCTTTTGAACGGCGGCGTGAACGCGACTGAA 3698  
Qy 2988 CGGACGCGACTACAGGTAAACGGGCGGCTTTACGGGCGGACTTGACAGCAACCGGCAAGAC 3047  
Db 3699 CGGACGCGACTACAGGTAAACGGGCGGCTTTACGGGCGGACTTGACAGCAACCGGCAAGAC 3758  
Qy 3048 GGGGCGAGCAATATGCCGACACCCCGCTGTTGCCGCTCTGGCGCGGATGTCGAATT 3107  
Db 3759 GGGGCGAGCAATATGCCGACACCCCGCTGTTGCCGCGCTTGGGCGCGGATGTCGAATT 3818



Qy 1668 TAGCCTGAACAGCAGCGCATTTGCTATCTGGCAGATACCGACCGATCCGGCGCAACGA 1727  
Db 2379 CAGCCTGAACAGCAGCGCATTTGCTATCTGGCAGATACCGACCAATCCGGCGCAACGA 2438  
Qy 1728 AACCGTGACATCAAGCGGATCTGCAGCTGGGGGGGAAAGGTAGCGTGTACACAGTTT 1787  
Db 2439 AACCGTACACATCAAGCGAGCTCTGCAGCTGGACGGCAAGAGTACGCTGTACACAGTTT 2498  
Qy 1788 GGGCAAACTGCTGAAGTGGAGGTPACGGCGATACCGCGCGCAAGCTGTACATCTCGGC 1847  
Db 2499 GGGCAAACTGCTGAAGTGGAGGTPACGGCGATATCCGGCGCAAGCTGTACATCTCGGC 2558  
Qy 1848 ACGGGCAAAAGGGGAGGCTATCTCAACAGTACCGAGCAACAGTGTTCCTTCAGTGC 1907  
Db 2559 ACGGGCAAAAGGGGAGGCTATCTCAACAGTACCGAGCAGCTGTTCCTTCAGTGC 2618  
Qy 1908 CGCCAAATCGGGGGGATTTCTTTCTTCAAAACATCGAAACCGAGCGTCTGCT 1967  
Db 2619 CGCCAAATCGGGGAGGATTTCTTTCTTCAAAACATCGAAACCGAGCGTCTGCT 2678  
Qy 1968 GGTCTCCCTCGACAGCTCGAAAAACAGCGGCGAGTGAAGGCGACAGCTGTCTATT 2027  
Db 2679 GGTCTCCCTCGACAGCTCGAAAAACAGCGGCGAGTGAAGGCGACAGCTGTCTATT 2738  
Qy 2028 TGTCCGTCCGGGCAATCGGCACAGGACTCTTCGGCAGCGGCACATTCGGCGCCCGCGG 2087  
Db 2739 TGTCCGTCCGGGCAATCGGCACAGGACTCTTCGGCAGCGGCACATTCGGCGCCCGCGG 2798  
Qy 2088 TCTGAACACGCGGTAGAACAGGGGCGGAGCAATCTGAAACCTGATGGTTCGAATCGA 2147  
Db 2799 TCTGAACACGCGGTAGAACAGGGGCGGAGCAATCTGAAACCTGATGGTTCGAATCGA 2858  
Qy 2148 TGCTCTCCGAATCATCCGCAACACCGGAGACGCTTGAACACTGCGGCGCGGACACAGA 2207  
Db 2859 TGCTCTCCGAATCATCCGCAACACCGGAGACGCTTGAACACTGCGGCGCGGACACAGA 2918  
Qy 2208 TATGCCGGGCAATCCGCCCTTACGGGCAACTTTCGGCAGCGGCGAGCGGTACAGCATGC 2267  
Db 2919 TATGCCGGGCAATCCGCCCTTACGGGCAACTTTCGGCAGCGGCGAGCGGTACAGCATGC 2978  
Qy 2268 GAATGCCGCGGACGCTGTACGCAATCTTCAACAGTCTCGCGGTACCGTCTATGCCGACAG 2327  
Db 2979 GAATGCCGCGGACGCTGTACGCAATCTTCAACAGTCTCGCGGTACCGTCTATGCCGACAG 3038  
Qy 2328 TACCGCGCCCATCGGATATGACAGGAGCGCGGTGAAAGCGGTATCGAGCGGTGGA 2387  
Db 3039 TACCGCGCCCATCGGATATGACAGGAGCGCGGTGAAAGCGGTATCGAGCGGTGGA 3098  
Qy 2388 CCACAACTGCTAGGGTCTGCGGTCTATCGCGAAACCCACAGAGCGGTGGAACGTGGGA 2447  
Db 3099 CCACAACTGCTAGGGTCTGCGGTCTATCGCGAAACCCACAGAGCGGTGGAACGTGGGA 3158  
Qy 2448 ACAGGCGGTGTTGAAGCAAAATGCGCGGAGTACCCAAACCGTCCGATTCGCCGCA 2507  
Db 3159 ACAGGCGGTGTTGAAGCAAAATGCGCGGAGTACCCAAACCGTCCGATTCGCCGCA 3218  
Qy 2508 AACCGGCAAAATACGACAGCGCGCCACACTGGGATGGACACACATGGAGCGA 2567  
Db 3219 AACCGGCAAAATACGACAGCGCGCCACACTGGGATGGGACGACACATGGAGCGA 3278  
Qy 2568 AAACAGTCAAAATGCAAAACCGACAGCATTTAGTCTGTTTGCAGGCATACGGCAGTGC 2627  
Db 3279 AAACAGTCAAAATGCAAAACCGACAGCATTTAGTCTGTTTGCAGGCATACGGCAGTGC 3338  
Qy 2628 GGGGATATCGGCTATCTCAAGGCGCTTCTCTACGGAGCGCTACAAACACAGCATCAG 2687  
Db 3339 GGGGATATCGGCTATCTCAAGGCGCTTCTCTACGGAGCGCTACAAACACAGCATCAG 3398  
Qy 2688 CGCAGCACCGGTGGGAGCAACATGCGGAAGGAGCGGTCAACGGCAGCGTGTATCGACT 2747  
Db 3399 CGCAGCACCGGTGGGAGCAACATGCGGAAGGAGCGGTCAACGGCAGCGTGTATCGACT 3458  
Qy 2748 GGGCGCACTGGCGGTGTCAACGTTTCCGTTTCCGCGAAACGGGAGATTTGACGGTGAAG 2807

Db 3459 GGGCGCACTGGCGGTGTCAACGTTCCGTTTCCGCAACGGAGATTTGACGGTCAAGG 3518  
Qy 2808 CGGTCTGGCTTACGACCTGCTCAACACAGGATGCATTGCCCAAAAAGCAGTGTCTTGGG 2867  
Db 3519 CGGTCTGGCTTACGACCTGCTCAACACAGGATGCATTGCCCAAAAAGCAGTGTCTTGGG 3578  
Qy 2868 CTGAGCGGCAACAGCTCACTGAAGGCACACTGCTCGGACTCGCGGCTCTGAAGCTGTC 2927  
Db 3579 CTGAGCGGCAACAGCTCACTGAAGGCACACTGCTCGGACTCGCGGCTCTGAAGCTGTC 3638  
Qy 2928 GCAACCCCTTGAGCGATAAAGCCGCTCTGTTTGAACGGCGGCGTGGAAACGCGCTGAA 2987  
Db 3639 GCAACCCCTTGAGCGATAAAGCCGCTCTGTTTGAACGGCGGCGTGGAAACGCGCTGAA 3698  
Qy 2988 CGGACGCGACTTACAGGTAACGGGCGGCTTTACGGCGGCGACTGCAGCAACGGCAAGAC 3047  
Db 3699 CGGACGCGACTTACAGGTAACGGGCGGCTTTACGGCGGCGACTGCAGCAACGGCAAGAC 3758  
Qy 3048 GGGGCGACGCAATATGCCGCACACCCGCTGTTGCCGCTTGGCGCGGATGTGCAAT 3107  
Db 3759 GGGGCGACGCAATATGCCGCACACCCGCTGTTGCCGCTTGGCGCGGATGTGCAAT 3818  
Qy 3108 CGGCAACGGCTGGAACGGCTTGGCACGTTTACAGCTACGCCGCTTCCAAACAGTACGGCAA 3167  
Db 3819 CGGCAACGGCTGGAACGGCTTGGCACGTTTACAGCTACGCCGCTTCCAAACAGTACGGCAA 3878  
Qy 3168 CCACAGCGGAGGCTGCGGCTAGGCTACCGGTTCC 3201  
Db 3879 CCACAGCGGAGGCTGCGGCTAGGCTACCGGTTCC 3912

RESULT 13  
AX236463 LOCUS AX236463 4170 bp DNA linear PAT 26-SEP-2001  
DEFINITION Sequence 156 from Patent WO0164922.  
ACCESSION AX236463  
VERSION AX236463.1 GI:15796053  
KEYWORDS synthetic construct.  
SOURCE ORGANISM synthetic construct.  
REFERENCE 1 (bases 1 to 4170)  
AUTHORS Arico, M.B., Comanducci, M.C., Galeotti, C.C., Masignani, V.C.,  
Guillani, M.M. and Pizzi, M.C.  
TITLE Heterologous expression of neisserial proteins  
JOURNAL Patent: WO 0164922-A 156 07-SEP-2001;  
Chiron Spa (IT)  
FEATURES  
source Location/Qualifiers  
1. 4170  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="961c-983"  
BASE COUNT 1185 a 1138 c 1119 g 728 t  
ORIGIN

Query Match 93.1%; Score 2983.2; DB 6; Length 4170;  
Best Local Similarity 97.1%; Pred. No. 0;  
Matches 3064; Conservative 0; Mismatches 48; Indels 42; Gaps 1;

Qy 90 CGGCGGGCGGCGCACTTCTGCGCCGCACTTCAATGACGGCGGCAACCGGTATCGGAGCAA 149  
Db 990 CGGCGGAGGCGGCACTTCTGCGCCGCACTTCAATGACGGCGGTACCGGTATCGGAGCAA 1049  
Qy 150 CAGCAGACACACAGCGAAATCAGCAGCAGTATCTTACCGCGGTATCAGAACGAAT 209  
Db 1050 CAGCAGACACACAGCGAAATCAGCAGCAGTATCTTACCGCGGTATCAGAACGAAT 1109  
Qy 210 GTGCAAAACAGACAGCAGTCTGTGCGCGGTGGGATGACGTTGGGGTTTACAGACAGGGA 269  
Db 1110 GTGCAAAACAGACAGCAGTCTGTGCGCGGTGGGATGACGTTGGGGTTTACAGACAGGGA 1169  
Qy 270 TGCCAAATCAATGCCCGCCCCCGGAAATCTGCATACCGGAGACTTTTACAAACCCCAATGA 329

Db	2250	CGGTACCACAGTGTGCTGACGACGCGCTCAGGACATFCGGTGCAGTGCAGTGCAGCGGTGGACAGCAAGTT	2309
Qy	1368	CGGTGGGGACTGCTGATCGGGTAAGGCCATTAAGAGGACCGCGCTCTTTCCGTTCCG	1427
Db	2310	CGGCTGGGACTGCTGGATCGGGTAAGGCCATTAAGAGGACCGCGCTCTTTCCGTTCCG	2369
Qy	1428	CGACTTTACGCGGATACGAAAGGTACATCCGATATGCGTACTCCTTCCGTAAAGCAT	1487
Db	2370	CGACTTTACGCGGATACGAAAGGTACATCCGATATGCGTACTCCTTCCGTAAAGCAT	2429
Qy	1488	TTACGGCACGGCGGCTGATCAAAAAGGCGCAGCCAACTGCAACTGCACGGCAACAA	1547
Db	2430	TTACGGCACGGCGGCTGATCAAAAAGGCGCAGCCAACTGCAACTGCACGGCAACAA	2489
Qy	1548	CACCTATACGGCAAAACCAATTATCGAAGCGGTTCCGTGTTGTACGGCAACAA	1607
Db	2490	CACCTATACGGCAAAACCAATTATCGAAGCGGTTCCGTGTTGTACGGCAACAA	2549
Qy	1608	ATCGGATATCGGCTGAAACCAAGGTGCGTGATTTATAACGGGGGGCATCCGGCGG	1667
Db	2550	ATCGGATATCGGCTGAAACCAAGGTGCGTGATTTATAACGGGGGGCATCCGGCGG	2609
Qy	1668	TAGCCTAACAGCGACGGCATTTCTCTATCTGGCAGATACCGACCCATCCGGCGCAACGA	1727
Db	2610	CAGCCTGAACAGCGACGGCATTTCTCTATCTGGCAGATACCGACCCATCCGGCGCAACGA	2669
Qy	1728	AACCGTCACATCAAGGGCATCTGCAGCTGGCGGCAAGGTACGCTGTACACAGTTT	1787
Db	2670	AACCGTCACATCAAGGGCATCTGCAGCTGGCGGCAAGGTACGCTGTACACAGTTT	2729
Qy	1788	GGGCAAACTGCTGAAAGTGGACGGTACGGCGATACGGCGCAAGCTGTACATGTCCGC	1847
Db	2730	GGGCAAACTGCTGAAAGTGGACGGTACGGCGATATTCGGCGCAAGCTGTACATGTCCGC	2789
Qy	1848	ACGGCGCAAGGGCAGGCTATCTCAACCGTACCGGACAACTGTTTCCCTTCTGAGTGC	1907
Db	2790	ACGGCGCAAGGGGCAAGCTATCTCAACAGTACCGGACAACTGTTTCCCTTCTGAGTGC	2849
Qy	1908	CGCCAAATCGGGCGGATTTCTTTCTTCAAAATCGAAACATCGAACCGGTTGCTGCT	1967
Db	2850	CGCCAAATCGGGCAGGATTTCTTTCTTCAAAATCGAAACATCGAACCGGCGCTGCT	2909
Qy	1968	GGCTTCCCTCGACAGGCTGAAAAACAGCGGCGAGTGAAGCGACAGCGTGTCTATT	2027
Db	2910	GGCTTCCCTCGACAGGCTGAAAAACAGCGGCGAGTGAAGCGACAGCGTGTCTATT	2069
Qy	2028	TGTCGGTCCGGCAATCCGCGCAGGACTGCTTCGGCAGCGGCACATTCGGCGCCCGCGG	2087
Db	2970	TGTCGGTCCGGCAATTCGCGCAGGACTGCTTCGGCAGCGGCACATTCGGCGCCCGCGG	3029
Qy	2088	TCTGAAACACGGCTAGAACAGGGCGGCAATCTGAAAACTTGATGGTTCGAACATGGA	2147
Db	3030	TCTGAAACACGGCTAGAACAGGGCGGCAATCTGAAAACTTGATGGTTCGAACATGGA	3089
Qy	2148	TGCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTTCGGCGCCCGCACAG	2207
Db	3090	TGCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTTCGGCGCCCGCACAG	3149
Qy	2208	TATCGCGGGCATTCGGCCCTACGGCGCACTTTCGGCGCAGCGGACGGTACAGCATGC	2267
Db	3150	TATCGCGGGCATTCGGCCCTACGGCGCACTTTCGGCGCAGCGGACGGTACAGCATGC	3209
Qy	2268	GAATGCGCGCAGGTTGATCGCATCTTCAACAGTCTCGCGCTACCGTCTATGCCGACAG	2327
Db	3210	GAATGCGCGCAGGTTGATCGCATCTTCAACAGTCTCGCGCTACCGTCTATGCCGACAG	3269
Qy	2328	TACGCGCGCCATTCGGATATCAGGGAGCGCGGTTGAAAGCGGTATCGGACGGTTGGA	2387
Db	3270	TACGCGCGCCATTCGGATATCAGGGAGCGCGGTTGAAAGCGGTATCGGACGGTTGGA	3329
Qy	2388	CCACACGCTACGGGTTGCGGCTCATTCGCGCAACCCCAACAGGACGGTGGACGTGGGA	2447
Db	3330	CCACACGCTACGGGTTGCGGCTCATTCGCGCAACCCCAACAGGACGGTGGACGTGGGA	3389



QY 2448 ACAGGCGGTGTTGAAGGCAAAATGCGCGCAGTACCCAAACCGTGGCATTTGCCGGA 2507  
Db 3390 ACAGGCGGTGTTGAAGGCAAAATGCGCGCAGTACCCAAACCGTGGCATTTGCCGGA 3449  
QY 2508 AACCGGCAAAATACACAGCAGCCGCCACACTGGGCATGGGACACAGCATGGAGCA 2567  
Db 3450 AACCGGCAAAATACACAGCAGCCGCCACACTGGGCATGGGACACAGCATGGAGCA 3509  
QY 2568 AAACAGTGAATGCAAAACACGACAGCATTAGTCTGTTTGGAGGCATACGCAACGATGC 2627  
Db 3510 AAACAGTGAATGCAAAACACGACAGCATTAGTCTGTTTGGAGGCATACGCAACGATGC 3569  
QY 2628 GGGCGATATCGGCTATCTCAAAGGCTGTCTCTACGAGCCTCAAAAACAGCATCAG 2687  
Db 3570 GGGCGATATCGGCTATCTCAAAGGCTGTCTCTACGAGCCTCAAAAACAGCATCAG 3629  
QY 2688 CGCAGCAGCGGTGGGAGCAATCGGAAAGCAGCGTCAAGGACACGCTGATGAGCT 2747  
Db 3630 CGCAGCAGCGGTGGGAGCAATCGGAAAGCAGCGTCAAGGACACGCTGATGAGCT 3689  
QY 2748 GGGCGACTGGGCGGTGTAAGCTTCCGTTTGCCGAACGGAGATTTGACGCTCGAAGG 2807  
Db 3690 GGGCGACTGGGCGGTGTAAGCTTCCGTTTGCCGAACGGAGATTTGACGCTCGAAGG 3749  
QY 2808 CGGTCTGCGCTACGACCTGCTCAACAGAGTATTCGCGCGAAAGGCGAGTCTTTGGG 2867  
Db 3750 CGGTCTGCGCTACGACCTGCTCAACAGAGTATTCGCGCGAAAGGCGAGTCTTTGGG 3809  
QY 2868 CTGGAGCGGCAACAGCCTCACTGAAGGACACTGTGTCGAGCTCGCGGTCTGAAGCTGTC 2927  
Db 3810 CTGGAGCGGCAACAGCCTCACTGAAGGACACTGTGTCGAGCTCGCGGTCTGAAGCTGTC 3869  
QY 2928 GCACACCTTTGAGCGATAAGCCGTCGTCTGTCACGCGGGGCTGGAACGCACTGAA 2987  
Db 3870 GCACACCTTTGAGCGATAAGCCGTCGTCTGTCACGCGGGGCTGGAACGCACTGAA 3929  
QY 2988 CGGACGCGACTACACGCTAACCGGCGGCTTTACCGGCGGAGTGTGACGCAACCGGCAAGAC 3047  
Db 3930 CGGACGCGACTACACGCTAACCGGCGGCTTTACCGGCGGAGTGTGACGCAACCGGCAAGAC 3989  
QY 3048 GGGGCGACGATATCCGACACCCGCTGTTGCGCGTCTGGGCGCGGATGTGCAATTT 3107  
Db 3990 GGGGCGACGATATCCGACACCCGCTGTTGCGCGTCTGGGCGCGGATGTGCAATTT 4049  
QY 3108 CGGCAACGCTGAAGCGCTTGGCAGCTTACAGCTACGCGCGGTTCCAAACAGTACGCGCA 3167  
Db 4050 CGGCAACGCTGAAGCGCTTGGCAGCTTACAGCTACGCGCGGTTCCAAACAGTACGCGCA 4109  
QY 3168 CCACAGCGGACGAGTGGCGGTAGGCTACCGGTTTC 3201  
Db 4110 CCACAGCGGACGAGTGGCGGTAGGCTACCGGTTTC 4143

RESULT 14  
LOCUS AX239749 4170 bp DNA linear PAT 26-SEP-2001  
DEFINITION Sequence 51 from Patent WO0164920.  
ACCESSION AX239749  
VERSION AX239749.1 GI:15797385  
KEYWORDS  
SOURCE synthetic construct.  
ORGANISM synthetic construct.  
REFERENCE 1 (bases 1 to 4170)  
AUTHORS Arico,M.B., Comanducci,M., Galeotti,C., Massignani,V., Giuliani,M.M.  
and Piza,M.  
TITLE Hybrid expression of neisserial proteins  
JOURNAL Patent: WO 0164920-A 51 07-SEP-2001;  
Chiron Spa (IT)  
FEATURES  
source Location/Qualifiers  
1..4170  
/organism="synthetic construct"

/db\_xref="taxon:32630"  
/note="961c-983"  
BASE COUNT 1185 a 1138 c 1119 g 728 t  
ORIGIN

Query Match 93.1%; Score 2983.2; DB 6; Length 4170;  
Best Local Similarity 97.1%; Pred. No. 0;  
Matches 3064; Conservative 0; Mismatches 48; Indels 42; Gaps 1;  
QY 90 CGGCGGCGGCGCACTTCTGCGCCGCACTTCAATGCGAGCGCACCGGTATCGGAGCAA 149  
Db 990 CGGCGGAGCGGCACTTCTGCGCCGCACTTCAATGCGAGCGGTATCGGAGCAA 1049  
QY 150 CAGCAGAGCAACAACAGCGAAATCAGCAGCATATCTTACGCCGGTATCAAGAAGCAAT 209  
Db 1050 CAGCAGAGCAACAACAGCGAAATCAGCAGCATATCTTACGCCGGTATCAAGAAGCAAT 1109  
QY 210 GTGCAAGACAGAAAGATGCTCTGTGCCGGTGGGATGAGCTTGGGTTTACAGACAGGA 269  
Db 1110 GTGCAAGACAGAAAGATGCTCTGTGCCGGTGGGATGAGCTTGGGTTTACAGACAGGA 1169  
QY 270 TGCCAAATCAATGCCCGCCCGCAATCTGCATACCGGAGACTTTACAAACCCAAATGA 329  
Db 1170 TGCCAAATCAATGCCCGCCCGCAATCTGCATACCGGAGACTTTTCCAAACCCAAATGA 1229  
QY 330 CGCATACAGAATTTGATCAACCTCAAACTGCAATTTGAAGCAGGCTATACAGGACGCG 389  
Db 1230 CGCATACAGAATTTGATCAACCTCAAACTGCAATTTGAAGCAGGCTATACAGGACGCG 1289  
QY 390 GGTAGAGTAGTATGCTGATACAGGGGAATCCGTCGCGACGATATCCTTTCCCGAAT 449  
Db 1290 GGTAGAGTAGTATGCTGATACAGGGGAATCCGTCGCGACGATATCCTTTCCCGAAT 1349  
QY 450 GTATGCGCAAGAGACAGCGCTATAACGAAATTTACAAAACCTATACGCGCTATATGG 509  
Db 1350 GTATGCGCAAGAGACAGCGCTATAACGAAATTTACAAAACCTATACGCGCTATATGG 1409  
QY 510 GAAGGAAGCGCTGAAGCAGGCGGTTAAAGACATTTAAAGCTTCTTTTCGACGATGAGC 569  
Db 1410 GAAGGAAGCGCTGAAGCAGGCGGTTAAAGACATTTAAAGCTTCTTTTCGACGATGAGC 1469  
QY 570 CGTTATAGAGACTGAAGCAACCGCAGGATATCCGCCACGTTAAAGAATCGGACAT 629  
Db 1470 CGTTATAGAGACTGAAGCAACCGCAGGATATCCGCCACGTTAAAGAATCGGACAT 1529  
QY 630 CGATGTGCTCCCATATTTATTTGGCGGGCTTCCGTTGGAGCGCAGACCTGCGAGGGTAT 689  
Db 1530 CGATTTGGTCTCCCATATTTATTTGGCGGGCTTCCGTTGGAGCGCAGACCTGCGAGGGTAT 1589  
QY 690 TCGCCCGGATGCGCGCTACACATAATGAATACGATGATGGAACCAAGCAAGCAATAT 749  
Db 1590 TCGCCCGGATGCGCGCTACACATAATGAATGATGATGGAACCAAGCAAGCAATAT 1649  
QY 750 GTCTGAGCCATCCGCAATGATGGTCAAGCTGGCGGAACGCTGGGTCGCGATCGTCAA 809  
Db 1650 GTCTGAGCCATCCGCAATGATGGTCAAGCTGGCGGAACGCTGGGTCGCGATCGTCAA 1709  
QY 810 TAACAGTTTGGCAACAACATCGAGGCGGCACTGCGGACCATTTCCAAATAGCCAAATTC 869  
Db 1710 TAACAGTTTGGCAACAACATCGAGGCGGCACTGCGGACCATTTCCAAATAGCCAAATTC 1769  
QY 870 GGAGGAGCAGTACCGCAAGCGTTGCTCCGCTATTTCCGCGGTGATAAACACAGACGAGG 929  
Db 1770 GGAGGAGCAGTACCGCAAGCGTTGCTCCGCTATTTCCGCGGTGATAAACACAGACGAGG 1829  
QY 930 TATCCGCTGATGCAACAGAGGATTTACGGCACTTGTCTTACCACATCCGTAATAAAAA 989  
Db 1830 TATCCGCTGATGCAACAGAGGATTTACGGCACTTGTCTTACCACATCCGTAATAAAAA 1889  
QY 990 CATGCTTTTCATTTTTCGCAAGCAATCACGCAAGCTCAGCCCAACACACTGACCCCT 1049  
Db 1890 CATGCTTTTCATTTTTCGCAAGCAATCACGCAAGCTCAGCCCAACACACTGACCCCT 1949



Qy	1050	ATTGCCATTTTATGAAAAAGATGCTCAAAAAGGCATTTATCACAGTCGACGCGGTAGACCG	1100
Db	1950	ATTGCCATTTTATGAAAAAGAGCGTCAAAAAGGCATTTATCACAGTCGACGCGGTAGACCG	2009
Qy	1110	CAGTGGAGAAAGTTCA-----A	1127
Db	2010	CAGTGGAGAAAGTTCAACGGGAATGTATGGAGAACCGGGTACAGAACCGCTTTGAGTA	2069
Qy	1128	TGGCTCCAACCATTTGGCGGAATTACTGCCATGTGGTGCCTATTCGGCACCCCTATGAAGCAAG	1187
Db	2070	TGGCTCCAACCATTTGGCGGAATTACTGCCATGTGGTGCCTTCGGCACCCCTATGAAGCAAG	2129
Qy	1188	CGTCCGTTTCACCCGGTACAAACCCGATTCAAATTCGCCGGAACATCTTTTCCGACCCCAT	1247
Db	2130	CGTCCGTTTCAACCGGTACAAACCCGATTCAAATTCGCCGGAACATCTTTTCCGACCCCAT	2189
Qy	1248	CGTAACCGGCACCGCGCTCTGCTGTCAGAAATACCCTGGATGACCAACGACACCT	1307
Db	2190	CGTAACCGGCACCGCGCTCTGCTGTCAGAAATACCCTGGATGACCAACGACACACCT	2249
Qy	1308	CGGTACCACGCTGCTGACAAACGGCTCAGGACATCGGTGCAGTCGGCGGTGACACGACAAATT	1367
Db	2250	CGGTACCACGTTGCTGACGACGGCTCAGGACATCGGTGCAGTCGGCGGTGACACGACAAATT	2309
Qy	1368	CGGCTGGGGACTGCTGGATCGCGGTAAAGCCATGAACGAGACCCCGTCTTTTCGTTTCGG	1427
Db	2310	CGGCTGGGGACTGCTGGATCGCGGTAAAGCCATGAACGAGACCCCGTCTTTTCGTTTCGG	2369
Qy	1428	CGACTTTACCGCGCATACGAAAGGTACATCCCATATTTCCTTACTCTTCCTTCGTACGACAT	1487
Db	2370	CGACTTTACCGCGCATACGAAAGGTACATCCCATATTTCCTTACTCTTCCTTCGTACGACAT	2429
Qy	1488	TTACAGCACGGCGGCGCTGATCAAAAAGCGCGCAGCAACTGCAACTGCACGGCAACAA	1547
Db	2430	TTACAGCACGGCGGCGCTGATCAAAAAGCGCGCAGCAACTGCAACTGCACGGCAACAA	2489
Qy	1548	CACCTATACGGGCAAAACCAATTATCGAAGGCGGTTTCGCTGGTGTGTACGGCAACACAA	1607
Db	2490	CACCTATACGGGCAAAACCAATTATCGAAGGCGGTTTCGCTGGTGTGTACGGCAACACAA	2549
Qy	1608	ATCGGATATCGCGCTCGAAACCAAGGTGCGCTGATTTTATACGGGGCGGATCCGGCGG	1667
Db	2550	ATCGGATATCGCGCTCGAAACCAAGGTGCGCTGATTTTATACGGGGCGGATCCGGCGG	2609
Qy	1668	TAGCCTCAACAGCAGCGCATTTGCTATCTGGCAGATACCGACCGATCCGGCGCAAAACGA	1727
Db	2610	CAGCCTCAACAGCAGCGCATTTGCTATCTGGCAGATACCGACCAATCCGGCGCAAAACGA	2669
Qy	1728	AACGTCGCATCAAAAGCGCATCTGCAGCTGGCGGCGGAAGGTACGCTGTACACACGTTT	1787
Db	2670	AACGTCGCATCAAAAGCGCATCTGCAGCTGGCGGCAAGGTACGCTGTACACACGTTT	2729
Qy	1788	GGCAAACTCTGAAAGTGGACGGTACGGGATACCGCGCGGCAAGCTGTACATGTCGGC	1847
Db	2730	GGCAAACTCTGAAAGTGGACGGTACGGGATATCGCGCGGCAAGCTGTACATGTCGGC	2789
Qy	1848	ACGCGCAAAAGGGCAGGCTATCTCAACCGTACCGGCAACAGTGTCCCTTCCTGAGTGC	1907
Db	2790	ACGCGCAAGGGGCAGGCTATCTCAACAGTACCGGCAACAGTGTCCCTTCCTGAGTGC	2849
Qy	1908	CGCAAAATCGGGCGGGATTTTCTTTCTTCAAAACATFCGAAACCGACGCTGTCTGCT	1967
Db	2850	CGCAAAATCGGGCAGGATTTTCTTTCTTCAAAACATFCGAAACCGACGCTGTCTGCT	2909
Qy	1968	GGCTTCCCTCGACAGCGTCGAAAAACACGGGGCAGTGAAGCGACACGCTGTCTTATTA	2027
Db	2910	GGCTTCCCTCGACAGCGTCGAAAAACACGGGGCAGTGAAGCGACACGCTGTCTTATTA	2969
Qy	2028	TGTCGCTCGGGCAATTCGCGCAGGACTCTTTTCGCGACGGCACATTCGCGCCCGCGG	2087
Db	2970	TGTCGCTCGGGCAATTCGCGCAGGACTCTTTTCGCGACGGCACATTCGCGCCCGCGG	3029
Qy	2088	TCTGAAACACCGCGTAGAACACAGGGCGGACGAATCTGGAAAAACCTTGATGTGCAACTGGA	2147

Db	3030	TCTGAAACACCCCGTAGAACAGCGCGCAGCAATCTGAAAAACCTGATGTGCGAACCTGGA	3089	
Qy	2148	TGCTTCGGAATCATCCGCACACCCGAGACGGTTGAAACTGGGCGCCGACCGCACAGA	2207	
Db	3090	TGCTTCGGAATCATCCGCACACCCGAGACGGTTGAAACTGGGCGACCGACCGCACAGA	3149	
Qy	2208	TATCGCGGCATCCGCCCTACCGCGCAACTTTCGCGCGCAGCGCAGCCGTACAGCATGC	2267	
Db	3150	TATCGCGGCATCCGCCCTACCGCGCAACTTTCGCGCGCAGCGCAGCCGTACAGCATGC	3209	
Qy	2268	GAATGCGCGCAGCGGTACGCATCTTCAACAGTCTTCGCGGTACCGTCTATGCGGCACAG	2327	
Db	3210	GAATGCGCGCAGCGGTACGCATCTTCAACAGTCTTCGCGGTACCGTCTATGCGGCACAG	3269	
Qy	2328	TACGCGCGCCATCCGCATATCAGGACCGCGGCTGAAGCCGTATCCGACGGTGTGGA	2387	
Db	3270	TACGCGCGCCATCCGCATATCAGGACCGCGGCTGAAGCCGTATCCGACGGTGTGGA	3329	
Qy	2388	CCACAAGCTACGGGTCTGCGCTCATCGCGCAACCCAACAGGACGGTGAAGCTGGGA	2447	
Db	3330	CCACAAGCAGCGGTCTGCGCTCATCGCGCAACCCAACAGGACGGTGAAGCTGGGA	3389	
Qy	2448	ACAGGGCGGTGTGAAGGCAAAATGCGCGCAGTACCCAAACCGTGGCATTTGCCGGCAA	2507	
Db	3390	ACAGGGCGGTGTGAAGGCAAAATGCGCGCAGTACCCAAACCGTGGCATTTGCCGGCAA	3449	
Qy	2508	AACGCGCGAAATACGACAGCAGCCGCCACACTGGGCATGGGACACACGACATGGACGA	2567	
Db	3450	AACGCGCGAAATACGACAGCAGCCGCCACACTGGGCATGGGACACACGACATGGACGA	3509	
Qy	2568	AAACAGTGCAAAATGCAAAAAACCGACAGCATTAGTCTGTTGAGGCATACGGCACATGC	2627	
Db	3510	AAACAGTGCAAAATGCAAAAAACCGACAGCATTAGTCTGTTGAGGCATACGGCACATGC	3569	
Qy	2628	GGCGCATATCGGTATCTCAAAGGCGCTGTCTCTACGGACGCTCAACGGCAGCTGATGCAGCT	2687	
Db	3570	GGCGCATATCGGTATCTCAAAGGCGCTGTCTCTACGGACGCTCAACGGCAGCTGATGCAGCT	3629	
Qy	2688	CCGACAGCAGCGGTGCGGACAAATCGGAGGACGCTCAACGGCAGCTGATGCAGCT	2747	
Db	3630	CCGACAGCAGCGGTGCGGACAAATCGGAGGACGCTCAACGGCAGCTGATGCAGCT	3689	
Qy	2748	GGCGCAGCTGGGCGGTGTCAACGTTCCGTTTGC CGCACGGAGATTTGACGGTCAAGG	2807	
Db	3690	GGCGCAGCTGGGCGGTGTCAACGTTCCGTTTGC CGCACGGAGATTTGACGGTCAAGG	3749	
Qy	2808	CGGTCTCGGTACGACCTGTCAAACAGGATGCATTCGCCGCAAAAGCAGTGTCTTTGGG	2867	
Db	3750	CGGTCTCGGTACGACCTGTCAAACAGGATGCATTCGCCGCAAAAGCAGTGTCTTTGGG	3809	
Qy	2868	CTGAGCGGCAACAGCTCACTCAAGGACACTGGTCGGACTCGGGCTGTAAGCTGTC	2927	
Db	3810	CTGAGCGGCAACAGCTCACTCAAGGACACTGGTCGGACTCGGGCTGTAAGCTGTC	3869	
Qy	2928	GCAACCCCTTGAGCGATAAAGCCGCTCTGTTTGCACGGCGGCGGTGGAACGCACTGAA	2987	
Db	3870	GCAACCCCTTGAGCGATAAAGCCGCTCTGTTTGCACGGCGGCGGTGGAACGCACTGAA	3929	
Qy	2988	CGGACGCGACTACACGATAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGCAAGAC	3047	
Db	3930	CGGACGCGACTACACGATAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGCAAGAC	3989	
Qy	3048	GGGGCAGCAATATCGGCACACCCGCTGTTGCCGGTCTGGGCGCGGATGTGCAATT	3107	
Db	3990	GGGGCAGCAATATCGGCACACCCGCTGTTGCCGGTCTGGGCGCGGATGTGCAATT	4049	
Qy	3108	CGGCAACGGCTGGAACGGCTTGGCAGTTACAGCTACGCGCGGTTCCTCAACAGTAGTCGCAA	3167	
Db	4050	CGGCAACGGCTGGAACGGCTTGGCAGTTACAGCTACGCGCGGTTCCTCAACAGTAGTCGCAA	4109	
Qy	3168	CCACAGCGGACGAGTCGGCGGTAGGCTACCGGTTCT	3201	



QY	1801	AAAGTGACGGGTACGGCGATGACCGCGGCAAGCTGTACATGTCTGGGCACCGCGCAAGGG	1860
Db	1744	AAAGTGACGGGTACGGCGATATATCGCGGCAAGCTGTACATGTCTGGGCACCGCGCAAGGG	1803
QY	1861	GCAGGCTATCTCAACCGTACCGGACACAGTGTTCCTTCCTGTAGTCCGCCCAAAATCGGG	1920
Db	1804	GCAGGCTATCTCAACAGTACCGGACGAGTGTTCCTTCCTGTAGTCCGCCCAAAATCGGG	1863
QY	1921	CGGGATTATTCTTCTTCAAAAACATCGAAACGACGGTGTCTGCTGGCTTCCTTCGCAC	1980
Db	1864	CAGATTATTCTTCTTCAAAAACATCGAAACGACGGCGCTGTCTTCCTTCCTTCGCAC	1923
QY	1981	AGCGTCGAAAAAAGCGGGCAGTGAAGCGACACCGTGTCTTATTATGTCCGTTCGGGGC	2040
Db	1924	AGCGTCGAAAAAAGCGGGCAGTGAAGCGACACCGTGTCTTATTATGTTCCTTCGGGGC	1983
QY	2041	AATGGCGCAGGACTGTTTCGGCAGCGGCACATTCGCGCGCCGCGGTGTGAACAGGCC	2100
Db	1984	AATGGCGCAGGACTGTTTCGGCAGCGGCACATTCGCGCGCCGCGGTGTGAACAGGCC	2043
QY	2101	GTAGACAGGGCGCAGCAATCTGGAACCTTGATGTCTGAACCTGGATGCTCCGGAATCA	2160
Db	2044	GTAGACAGGGCGCAGCAATCTGGAACCTTGATGTCTGAACCTGGATGCTCCGGAATCA	2103
QY	2161	TCGCGAACACCCGAGAGCGTTGAAACTGCGCGCGCGACCGCACAGATATGCCGGGCATC	2220
Db	2104	TCGCGAACACCCGAGAGCGTTGAAACTGCGCGCGCGACCGCACAGATATGCCGGGCATC	2163
QY	2221	CGCCCTTACCGGCGCACTTTTCGCGCAGCGCAGCGGTACAGCATGCGAATGCCGCCGAC	2280
Db	2164	CGCCCTTACCGGCGCACTTTTCGCGCAGCGCAGCGGTACAGCATGCGAATGCCGCCGAC	2223
QY	2281	GGTGTAGCATCTTCAACAGTCTCGCGCTACCGTCTATGCCACAGTACCGCGGCCAT	2340
Db	2224	GGTGTAGCATCTTCAACAGTCTCGCGCTACCGTCTATGCCACAGTACCGCGGCCAT	2283
QY	2341	GCGGATATGAGGAGCGCGGCTGAAAGCGTATCGGACGGTGGACCAACAGCGTACG	2400
Db	2284	GCGGATATGAGGAGCGCGGCTGAAAGCGTATCGGACGGTGGACCAACAGCGCAGC	2343
QY	2401	GGTCTGGGCTCATCGGCAAAACCCAACAGGACGGTGGAAAGTGGGAACAGGCGGTGT	2460
Db	2344	GGTCTGGGCTCATCGGCAAAACCCAACAGGACGGTGGAAAGTGGGAACAGGCGGTGT	2403
QY	2461	GAAGGCAAAATGCGCGCAGTACCAACCGTTCGGCATTCGCCGGAACCCGCGGCAAT	2520
Db	2404	GAAGGCAAAATGCGCGCAGTACCAACCGTTCGGCATTCGCCGGAACCCGCGGCAAT	2463
QY	2521	ACGACAGCGCCCGCCACTGGGCATGGGACACAGCACATGGAGCGAAACAGTGCAAAT	2580
Db	2464	ACGACAGCGCCCGCCACTGGGCATGGGACCGACACATGGAGCGAAACAGTGCAAAT	2523
QY	2581	GCAAAACCGACAGCATTAGTCTGTTTTCAGGCGATACGGCAGCATCGGGCGGATACGGC	2640
Db	2524	GCAAAACCGACAGCATTAGTCTGTTTTCAGGCGATACGGCAGCATCGGGCGGATACGGC	2583
QY	2641	TATCTCAAAAGGCTGTCTCTACGACGCTACAAAACAGCATACCGCGCAGCACCGGT	2700
Db	2584	TATCTCAAAAGGCTGTCTCTACGACGCTACAAAACAGCATACCGCGCAGCACCGGT	2643
QY	2701	GCGGAGCAACATCGGGAAGCAGCGTCAACGGCACGCTGATGCAGTGGGCGCACTGGGC	2760
Db	2644	GCGGAGCAACATCGGGAAGCAGCGTCAACGGCACGCTGATGCAGTGGGCGCACTGGGC	2703
QY	2761	GGTGTCAAGCTTCCGTTTTCGCCCAACGGGAGATTTGACGGTTCGAAGCGGTCTGCGCTAC	2820
Db	2704	GGTGTCAAGCTTCCGTTTTCGCCCAACGGGAGATTTGACGGTTCGAAGCGGTCTGCGCTAC	2763
QY	2821	GACCTGCTCAACAGAGTGCATTCGCGCGGAAAAAGGAGGTGCTTTGGGCTGGAGCGGCAAC	2880
Db	2764	GACCTGCTCAACAGAGTGCATTCGCGCGGAAAAAGGAGGTGCTTTGGGCTGGAGCGGCAAC	2823
QY	2881	AGCCTCACTGAAGGCACACACTGGTCTGGACTCGCGGCTCTGAAGCTGTCTGCAACCCCTTGAGC	2940
Db	2824	AGCCTCACTGAAGGCACACTGGTCTGGACTCGCGGCTCTGAAGCTGTCTGCAACCCCTTGAGC	2883
QY	2941	GATAAAGCCCTCTCTGTTTGAACGGGGCGGTGGAACGGCAGCTGAACGACGCGACTAC	3000
Db	2884	GATAAAGCCCTCTCTGTTTGAACGGGGCGGTGGAACGGCAGCTGAACGACGCGACTAC	2943
QY	3001	ACGGTAAACGGGGCGCTTACCGGGCGGCTGACGACACCGGCAAGCGGGGGCACCAAT	3060
Db	2944	ACGGTAAACGGGGCGCTTACCGGGCGGCTGACGACACCGGCAAGCGGGGGCACCAAT	3003
QY	3061	ATCCGCACACCCCGCTGTTGCGGCTCTGGCGCGGATGTCGAATTCGGCAACGGCTGG	3120
Db	3004	ATCCGCACACCCCGCTGTTGCGGCTCTGGCGCGGATGTCGAATTCGGCAACGGCTGG	3063
QY	3121	AACGGCTTGGCAGCTTACAGCTACGGCGGTTCCAAACAGTACGGCAACACACGCGGACGA	3180
Db	3064	AACGGCTTGGCAGCTTACAGCTACGGCGGTTCCAAACAGTACGGCAACACACGCGGACGA	3123
QY	3181	GTCCGGCTAGGCTACCGGTTTC 3201	
Db	3124	GTCCGGCTAGGCTACCGGTTTC 3144	
RESULT 16			
AX239717			
LOCUS	AX239717	Sequence 19 from Patent WO0164920.	3939 bp DNA linear PAT 26-SEP-2001
DEFINITION	AX239717	Accession	
ACCESSION	AX239717	Version	
VERSION	AX239717.1	GI:15797370	
KEYWORDS			
SOURCE		synthetic construct.	
ORGANISM		artificial construct.	
		artificial sequences.	
REFERENCE	1 (bases 1 to 3939)		
AUTHORS	Arico, M. B., Comanducci, M., Galeotti, C., Masignani, V., Giuliani, M. M. and Pizza, M.		
TITLE	Hybrid expression of neisserial proteins		
JOURNAL	Patent: WO 0164920-A 19 07-SEP-2001;		
Chiron Spa (It)			
FEATURES		Location/Qualifiers	
source	1..3939	/organism="synthetic construct"	
		/db_xref="taxon:32630"	
		/note="delta983-741"	
BASE COUNT	1046 a 1100 c 1107 g 686 t		
ORIGIN			
Query Match 92.8%; Score 2971.8; DB 6; Length 3939;			
Best Local Similarity 97.2%; Pred. No. 0;			
Matches 3052; Conservative 0; Mismatches 47; Indels 42; Gaps 1;			
QY	103	ACTTCTCGCGCCGCTTCAATGACGGCGGCGGTATCGGCAGCAACAGCAGGCAACA	162
Db	4	ACTTCTCGCGCCGCTTCAATGACGGCGGTTACCGGTATCGGCAGCAACAGCAGGCAACA	63
QY	163	ACAGCGAAATCAGCAGCATGTTTACGGCGGTATCAAGAACGAAATGTGCAAGACAGA	222
Db	64	ACAGCGAAATCAGCAGCATGTTTACGGCGGTATCAAGAACGAAATGTGCAAGACAGA	123
QY	223	AGCATGCTCTGTGCGGCTCGGGATGACGTTGGCGTTACAGACAGGGATGCCAAATCAAT	282
Db	124	AGCATGCTCTGTGCGGCTCGGGATGACGTTGGCGTTACAGACAGGGATGCCAAATCAAT	183
QY	283	GCCCCCCCCCGGAATCTGCATACCGGAGACTTTACAAACCCAAATACGCATACAAAGAT	342
Db	184	GCCCCCCCCCGGAATCTGCATACCGGAGACTTTCCAAACCCAAATACGCATACAAAGAT	243
QY	343	TTGATCAACCTCAAACTGCAATTTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGT	402
Db	244	TTGATCAACCTCAAACTGCAATTTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGT	303

QY 403 ATCGTCGATACAGGCGAATCCGTCGGCAGCATATCCTTTCCGAACTGTATGGCAGAAAA 462  
Db 304 ATCGTCGACACAGGCGAATCCGTCGGCAGCATATCCTTTCCGAACTGTATGGCAGAAAA 363  
QY 463 GAAACAGGCTATAACGAAAAATTACAAAACTATACGGCGTATATCGGGAAGAGCGCT 522  
Db 364 GAAACAGGCTATAACGAAAAATTACAAAACTATACGGCGTATATCGGGAAGAGCGCT 423  
QY 523 GAAAGCAGGCGGTAAAGACATTAAGAGCTTCTTCGACGATGAGCGCGTTATAGAGACT 582  
Db 424 GAAAGCAGGCGGTAAAGACATTAAGAGCTTCTTCGACGATGAGCGCGTTATAGAGACT 483  
QY 583 GAAAGCAGGCGGTAAAGACATTAAGAGCTTCTTCGACGATGAGCGCGTTATAGAGACT 642  
Db 484 GAAAGCAGGCGGTAAAGACATTAAGAGCTTCTTCGACGATGAGCGCGTTATAGAGACT 543  
QY 643 CATATTTATGGCGGCGTTCGTCGACGCGAGCTGAGCGGTTATTCGCGCCGATCGG 702  
Db 544 CATATTTATGGCGGCGTTCGTCGACGCGAGCTGAGCGGTTATTCGCGCCGATCGG 603  
QY 703 AGGCTACACATAATGAATACGCATGATGAACCAAGAACGAAATATGTCTGACGCCATC 762  
Db 604 AGGCTACACATAATGAATACGCATGATGAACCAAGAACGAAATATGTCTGACGCCATC 663  
QY 763 CGCAATGTCATGGTCAAGCTGGCGGAACGTGGCGTGCATCGTCAATAAAGAGTTTGGGA 822  
Db 664 CGCAATGTCATGGTCAAGCTGGCGGAACGTGGCGTGCATCGTCAATAAAGAGTTTGGGA 723  
QY 823 ACAACATCGAGGCGAGGCACTGCGGCAACATTTCCAAATAGCCAAATTCGGAGGAGCAGTAC 882  
Db 724 ACAACATCGAGGCGAGGCACTGCGGCAACATTTCCAAATAGCCAAATTCGGAGGAGCAGTAC 783  
QY 883 CGCCAAAGGCTTCTCGCTTATTCGCGGCGTATATAACAGAGGAGGTTATCGGCTGATG 942  
Db 784 CGCCAAAGGCTTCTCGCTTATTCGCGGCGTATATAACAGAGGAGGTTATCGGCTGATG 843  
QY 943 CAACAGAGCGATTACGGCAACTTGTCTACCACTCCGTAATAAAGAGTGTTCATTT 1002  
Db 844 CAACAGAGCGATTACGGCAACTTGTCTACCACTCCGTAATAAAGAGTGTTCATTT 903  
QY 1003 TTTTCGGCAAGCAATGACGCGCAAGCTCAGCCCAACACACTGACCTTATGGCAATTTAT 1062  
Db 904 TTTTCGACAGGCAATGACGCGCAAGCTCAGCCCAACACACTGACCTTATGGCAATTTAT 963  
QY 1063 GAAAAAGATGCTCAAAAAGGCAATATACAGTCGCGAGCGTAGACCGAGTGGAGAAAAG 1122  
Db 964 GAAAAAGATGCTCAAAAAGGCAATATACAGTCGCGAGCGTAGACCGAGTGGAGAAAAG 1023  
QY 1123 TTCA-----ATGGCTCCCAACCAT 1140  
Db 1024 TTCAAAACGGGAAATGTATGGAGAACCGGTACAGAACCGCTTGAGTATGGCTCCCAACCAT 1083  
QY 1141 TCGGGAATTTACTGTCATGTGGTGCCTATTCGGCACCCCTATGAAGCAAGGTCGTTTCACC 1200  
Db 1084 TCGGGAATTTACTGTCATGTGGTGCCTATTCGGCACCCCTATGAAGCAAGGTCGTTTCACC 1143  
QY 1201 CGTACAAACCCGATTCAAATTCGCGGAACATCCTTTTCGCAACCCATGTAACCGGCAG 1260  
Db 1144 CGTACAAACCCGATTCAAATTCGCGGAACATCCTTTTCGCAACCCATGTAACCGGCAG 1203  
QY 1261 GCGGCTCTGCTCTGCAGAAATACCGGTGGATGAGCAAGCAGCAACTCGGTACCAAGCTG 1320  
Db 1204 GCGGCTCTGCTCTGCAGAAATACCGGTGGATGAGCAAGCAGCAACTCGGTACCAAGCTG 1263  
QY 1321 CTGACAACGCGCTCAGACATCGGTCAGTCGCGTGGCAGCAAGTTCGGCTGGGAGCTG 1380  
Db 1264 CTGACAACGCGCTCAGACATCGGTCAGTCGCGTGGCAGCAAGTTCGGCTGGGAGCTG 1323  
QY 1381 CTGAGTGGGCTAAGGCCATGAACGACCGCGTCTCTTCGCTTCGGGAGCTTTACCGCC 1440  
Db 1324 CTGAGTGGGCTAAGGCCATGAACGACCGCGTCTCTTCGCTTCGGGAGCTTTACCGCC 1383  
QY 1441 GATACGAAAGGTACATCCGATATTCGCTACTCTCTCCGTACGACATTTTCAGGCGAGGCG 1500

Db 1384 GATACGAAAGGTACATCCGATATTTGCCCTACTCTTCGCTAACGACATTTTCAGGACGCGGC 1443  
QY 1501 GGCCTGATCAAAAAAGGCGGAGCAACTGCAACTGCAAGCAACCAACACCTATATACGGGC 1560  
Db 1444 GGCCTGATCAAAAAAGGCGGAGCAACTGCAACTGCAAGCAACCAACACCTATATACGGGC 1503  
QY 1561 AAAACCAATTATCGAAGCGGTTTCGCTGGTGTGTACGCAACCAACATCGGATATGCGC 1620  
Db 1504 AAAACCAATTATCGAAGCGGTTTCGCTGGTGTGTACGCAACCAACATCGGATATGCGC 1563  
QY 1621 GTCGAAACCAAGGTGCGCTGATTTATACGCGGGGGGATCCGGGGTAGCTGTAACAGC 1680  
Db 1564 GTCGAAACCAAGGTGCGCTGATTTATACGCGGGGGGATCCGGGGTAGCTGTAACAGC 1623  
QY 1681 GACGCAATTTGCTATCTGCGAGATACCGACGATCCGCGCAAAAGCAACCTGTCACATC 1740  
Db 1624 GACGCAATTTGCTATCTGCGAGATACCGACCAATCCGGCGCAAAAGCAACCTGTCACATC 1683  
QY 1741 AAAGCGATCTGCAAGCTGGGCGGCAAGGTACGCTGTACACACGTTTGGGCAAACTGCTG 1800  
Db 1684 AAAGCGATCTGCAAGCTGGGCGGCAAGGTACGCTGTACACACGTTTGGGCAAACTGCTG 1743  
QY 1801 AAAGTGGACGCTAGCGGATGACGCGGCGCAAGCTGTACATGTTCGGCACGCGGCAAAAGG 1860  
Db 1744 AAAGTGGACGCTAGCGGATGACGCGGCGCAAGCTGTACATGTTCGGCACGCGGCAAAAGG 1803  
QY 1861 GCAGCTATCTCAACGTTACCGGACACGCTGTTCCTCTTCCTGAGTCCGCCCAAAATCCGG 1920  
Db 1804 GCAGCTATCTCAACGTTACCGGAGCAGCTGTTCCTCTTCCTGAGTCCGCCCAAAATCCGG 1863  
QY 1921 CGGATTTATCTTCTTTCACAAACATCGAAACCGAGCGTGTCTGCTGGCTTCCCTCGAC 1980  
Db 1864 CAGGATTTATCTTCTTTCACAAACATCGAAACCGAGCGGCTGTCTGGCTTCCCTCGAC 1923  
QY 1981 AGGCTCGAAAAACAGCGGCGAGTGAAGCGGACAGCTGTCTTATTTGCTGCTCGCGGC 2040  
Db 1924 AGCCTCGAAAAACAGCGGCGAGTGAAGCGGACAGCTGTCTTATTTGCTGCTCGCGGC 1983  
QY 2041 AATCGGCGACGAGCTGCTTCGCGAGCGGACATTCGCGCGCCCGCGGCTCTGAACACAGCC 2100  
Db 1984 AATCGGCGACGAGCTGCTTCGCGAGCGGACATTCGCGCGCCCGCGGCTCTGAACACAGCC 2043  
QY 2101 GTAGAACAGGCGGCGAGCAATCTCGAAAACTGTATGTCGAACCTGGATGCTCTCGCAATCA 2160  
Db 2044 GTAGAACAGGCGGCGAGCAATCTCGAAAACTGTATGTCGAACCTGGATGCTCTCGCAATCA 2103  
QY 2161 TCCGCAACACCGGAGAGCGGTTGAACCTGCGGCGCGGCGGCGGACAGATATGCGGCGGATC 2220  
Db 2104 TCCGCAACACCGGAGAGCGGTTGAACCTGCGGCGCGGCGGCGGACAGATATGCGGCGGATC 2163  
QY 2221 CGCCCTTACGGCGCAACTTTCCGCGCAGCGGCGGCGGCTACAGCATGCGCAATTCGCGCGGAC 2280  
Db 2164 CGCCCTTACGGCGCAACTTTCCGCGCAGCGGCGGCGGCTACAGCATGCGCAATTCGCGCGGAC 2223  
QY 2281 GGTGTACGCACTCTTCAACAGTCTCGCCGCTACCGCTTATGCGGACAGTACCGCGGCCAT 2340  
Db 2224 GGTGTACGCACTCTTCAACAGTCTCGCCGCTACCGCTTATGCGGACAGTACCGCGGCCAT 2283  
QY 2341 GCCGATATGAGGAGACGCGGCTGAAACCGCTATCGGACGGGTTGGACACACAGCTACG 2400  
Db 2284 GCCGATATGAGGAGACGCGGCTGAAACCGCTTATCGGACGGGTTGGACACACAGCTACG 2343  
QY 2401 GGTCTGCGCGTCATCGCGCAAAACCAACAGGACGCTGAAACCTGGGAAACGCGCGGCTGT 2460  
Db 2344 GGTCTGCGCGTCATCGCGCAAAACCAACAGGACGCTGAAACCTGGGAAACGCGCGGCTGT 2403  
QY 2461 GAAAGCAAAATTCGCGCGGAGTACCCAAACCGCTCGCATTTGCGCGCAAAACCGCGGCAAAAT 2520  
Db 2404 GAAAGCAAAATTCGCGCGGAGTACCCAAACCGCTCGCATTTGCGCGCAAAACCGCGGCAAAAT 2463  
QY 2521 AGGACAGGCGGCGCACACTTGGGCTATGGGACACAGCATATGGAGCGAAACAGTCCAAT 2580

Db 2464 ACGACAGCAGCCGCCACACTGGCGATGGGACGACACATGGAGCGAAAAACAGTGCAAAT 2523  
QY 2581 GCAAAAACGACAGCATTAGTCTGTTTGCAGGCGATACGGCAGCATACGGCGGATATCGGC 2640  
Db 2524 GCAAAAACGACAGCATTAGTCTGTTTGCAGGCGATACGGCAGCATACGGCGGATATCGGC 2583  
QY 2641 TATCTCAAAAGGCGTGTCTCTACGACGCTCAAAAAACAGCATACGCCGCGACACCGGT 2700  
Db 2584 TATCTCAAAAGGCGTGTCTCTACGACGCTCAAAAAACAGCATACGCCGCGACACCGGT 2643  
QY 2701 GCGGAGCATACCGGAAGCAGCGTCAACGGCACGCTGATGACGTGGCGGCGACCTGGG 2760  
Db 2644 GCGGAGCAACATCGGAAGCAGCGTCAACGGCACGCTGATGACGTGGCGGCGACCTGGG 2703  
QY 2761 GGTGTCAACGCTTCGCTTGGCGCAACGGGAGATTGACGGTCCGAAGCGGTCTGCGCTAC 2820  
Db 2704 GGTGTCAACGCTTCGCTTGGCGCAACGGGAGATTGACGGTCCGAAGCGGTCTGCGCTAC 2763  
QY 2821 GACCTGCTCAACAGGATGCTTCGCGGAAAAAGGAGTGTCTTGGCTGGAGCGCAAC 2880  
Db 2764 GACCTGCTCAACAGGATGCTTCGCGGAAAAAGGAGTGTCTTGGCTGGAGCGCAAC 2823  
QY 2881 AGCCTCAGTGAAGCAGCATGCTGGGACTCGCGGCTGTAAGCTGTCGCAACCTTTGAGC 2940  
Db 2824 AGCCTCAGTGAAGCAGCATGCTGGGACTCGCGGCTGTAAGCTGTCGCAACCTTTGAGC 2883  
QY 2941 GATAAAGCGTCTGTTTGAACGGCGGCGTGAACGGCAGCTGTAAGCTGTCGCAACCTTTGAGC 3000  
Db 2884 GATAAAGCGTCTGTTTGAACGGCGGCGTGAACGGCAGCTGTAAGCTGTCGCAACCTTTGAGC 2943  
QY 3001 ACGGTAACGGCGGCTTTACCGCGGCGACTGTCAGCAACCGGCAAGCGGGGCGACGCAAT 3060  
Db 2944 ACGGTAACGGCGGCTTTACCGCGGCGACTGTCAGCAACCGGCAAGCGGGGCGACGCAAT 3003  
QY 3061 ATCGCGACACCGCTGTTGCGGCTGTCGGGCGGCGATGCGAAATTCGGCAACGGCTGG 3120  
Db 3004 ATCGCGACACCGCTGTTGCGGCTGTCGGGCGGCGATGCGAAATTCGGCAACGGCTGG 3063  
QY 3121 AACGGCTTGGCAGTACAGCTACGGGTTCCAAACAGTACGGCACACCGGACGA 3180  
Db 3064 AACGGCTTGGCAGTACAGCTACGGGTTCCAAACAGTACGGCACACCGGACGA 3123  
QY 3181 GTCGGCTAGGCTACCGGTTTC 3201  
Db 3124 GTCGGCTAGGCTACCGGTTTC 3144

RESULT 17  
AX236413  
LOCUS AX236413 4179 bp DNA linear PAT 26-SEP-2001  
DEFINITION Sequence 106 from Patent WO0164922.  
ACCESSION AX236413  
VERSION AX236413.1 GI:15796029  
KEYWORDS synthetic construct.  
SOURCE synthetic construct  
ORGANISM artificial sequences.  
REFERENCE 1 (bases 1 to 4179)  
AUTHORS Arico,M.B., Comanducci,M.C., Galeotti,C.C., Massignani,V.C.,  
Guilianini,M.M. and Pizzi,M.C.  
TITLE Heterologous expression of neisserial proteins  
JOURNAL Patent: WO 0164922-A 106 07-SEP-2001;  
Chiron Spa (IT)  
FEATURES  
source Location/Qualifiers  
1..4179  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="deltaG983-961c"  
BASE COUNT 1187 a 1140 c 1121 g 731 t  
ORIGIN  
Query Match 92.8%; Score 2971.8; DB 6; Length 4179;  
Best Local Similarity 97.2%; Pred. No. 0;

Matches 3052; Conservative 0; Mismatches 47; Indels 42; Gaps 1;  
QY 103 ACTTCTGCGCCCGACATTCAATGCGGGCGCACCGGTATCGGAGCAACACGACGAGAACA 162  
Db 4 ACTTCTGCGCCCGACATTCAATGCGGGGTACCGGTATCGGAGCAACACGACGAGAACA 63  
QY 163 ACAGCGAAATACAGCAGCAGTATCTTACGCGCGGTATCAAGAACGAAATGTGCAAGACAGA 222  
Db 64 ACAGCGAAATACAGCAGCAGTATCTTACGCGCGGTATCAAGAACGAAATGTGCAAGACAGA 123  
QY 223 AGCATGCTGTGTCGCGGTTCGGGATGAGCTTTCGGGTATACAGACAGGGATGCCAAATCAAT 282  
Db 124 AGCATGCTGTGTCGCGGTTCGGGATGAGCTTTCGGGTATACAGACAGGGATGCCAAATCAAT 183  
QY 283 GCCCGCCCCCGCAATCTGCAATACCGGAGACTTTTACAAACCCAAATGACGATACAAAGAT 342  
Db 184 GCCCGCCCCCGCAATCTGCAATACCGGAGACTTTTCCAAACCCAAATGACGATACAAAGAT 243  
QY 343 TTGATCAACCTCAACCTTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGT 402  
Db 244 TTGATCAACCTCAACCTTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGT 303  
QY 403 ATCGTGTATACAGCGGAGTTCGGTTCGGGAGCATATCTTTCGCAACTGTATGGCAGAAAA 462  
Db 304 ATCGTGTATACAGCGGAGTTCGGTTCGGGAGCATATCTTTCGCAACTGTATGGCAGAAAA 363  
QY 463 GAACAGGCTATACGAAATTAACAAACTATACGCGGTATATGCGGAAGGAGCGCCT 522  
Db 364 GAACAGGCTATACGAAATTAACAAACTATACGCGGTATATGCGGAAGGAGCGCCT 423  
QY 523 GAAGAGCGGCGGTAAAGACATTTAAAGCTTCTTTCGACGATGAGGCGCTTATAGAGACT 582  
Db 424 GAAGAGCGGCGGTAAAGACATTTAAAGCTTCTTTCGACGATGAGGCGCTTATAGAGACT 483  
QY 583 GAAGCAAGCGGAGGATATCCGCCAGTAAAAGAAATTCGAGACACATCGATGTGGTCTCC 642  
Db 484 GAAGCAAGCGGAGGATATCCGCCAGTAAAAGAAATTCGAGACACATCGATTTTGGTCTCC 543  
QY 643 CATATTATTGCGGGCGTTCGGTGGAGCGGAGCTTCGAGCGGTATTCGCGCGGATGCG 702  
Db 544 CATATTATTGCGGGCGTTCGGTGGAGCGGAGCTTCGAGCGGTATTCGCGCGGATGCG 603  
QY 703 ACGCTACACATAATGAATACGATGATGAACCAAGAACGAAATTAATCTCTGCGGCCATC 762  
Db 604 ACGCTACACATAATGAATACGATGATGAACCAAGAACGAAATTAATCTCTGCGGCCATC 663  
QY 763 CGCAATGCTAGGCTCAAGCTGGGGAACGTGGCGTGGCGATCGTCATATACAGTTTGGGA 822  
Db 664 CGCAATGCTAGGCTCAAGCTGGGGAACGTGGCGTGGCGATCGTCATATACAGTTTGGGA 723  
QY 823 ACAACATCGAGGCGAGCAGCTGCGGACCATTTCCAAATAGCCAATTCGGAGGACGATAC 882  
Db 724 ACAACATCGAGGCGAGCAGCTGCGGACCATTTCCAAATAGCCAATTCGGAGGACGATAC 783  
QY 883 CGCAAGCGTGTCTGCGCTATTCGGGCGGTGATAAAACAGACAGGATTCGCGCTGATG 942  
Db 784 CGCAAGCGTGTCTGCGCTATTCGGGCGGTGATAAAACAGACAGGATTCGCGCTGATG 843  
QY 943 CAACAGAGCATATACGCAACTTGTCTACACATCGGTATATAAAACATGCTTTTCATT 1002  
Db 844 CAACAGAGCATATACGCAACTTGTCTACACATCGGTATATAAAACATGCTTTTCATT 903  
QY 1003 TTTTTCGCAAGCATGACGACAGCTTCAGCCCAACACACTGACCCCTATTTCGCCATTTAT 1062  
Db 904 TTTTTCGCAAGCATGACGACAGCTTCAGCCCAACACACTGACCCCTATTTCGCCATTTAT 963  
QY 1063 GAAAAAGATGCTCAAAAAGCATTTATCACAGTTCGAGCGGTAGACCAGTGGAGAAAAAG 1122  
Db 964 GAAAAAGATGCTCAAAAAGCATTTATCACAGTTCGAGCGGTAGACCAGTGGAGAAAAAG 1023  
QY 1123 TTCA-----ATGGCTCCCAACCAT 1140  
Db 1024 TTCAACGGGAAATGTATGGAGAACCGGGTACAGAACCGGCTTGAGTATGGCTCCCAACCAT 1083

Qy	1141	TGCGGAATTACTGCCATGTGGTGCCCTATTCGGCACCCCTATGAAGCAAGGTCCTGTTTCACC	1200
Db	1084	TGCGGAATTACTGCCATGTGGTGCCCTGTGGCACCCCTATGAAGCAAGCGTCGCTTTCAAC	1143
Qy	1201	CGTACAAACCCGATTCAAATTCGGGACATCTTTTCCGCACCCATCGTAAACCGGCACG	1260
Db	1144	CGTACAAACCCGATTCAAATTCGGGACATCTTTTCCGCACCCATCGTAAACCGGCACG	1203
Qy	1261	GCGGCTCTGCTGTCAGAAATACCCGTGGATGAGCAACAGCAACCTGCGTACCACGCTG	1320
Db	1204	CGGGCTCTGCTGTCAGAAATACCCGTGGATGAGCAACAGCAACCTGCGTACCACGCTG	1263
Qy	1321	CTGACAACGGCTCAGGACATCGGTGAGTTCGGGTGGACGCAAGTTTCGGTGGGGACTG	1380
Db	1264	CTGACGAGGCTCAGGACATCGGTGAGTTCGGGTGGACGCAAGTTTCGGTGGGGACTG	1323
Qy	1381	CTGGATCGGGTAAGGCCATGAACGGACCCGCTCCCTTCGGTTCGGGACTTTACGCC	1440
Db	1324	CTGGATCGGGTAAGGCCATGAACGGACCCGCTCCCTTCGGTTCGGGACTTTACGCC	1383
Qy	1441	GATACGAAAGGTACATCCGATATTGCTACTCTTCCGTAAAGCATTTTCAGGCACGGCC	1500
Db	1384	GATACGAAAGGTACATCCGATATTGCTACTCTTCCGTAAAGCATTTTCAGGCACGGCC	1443
Qy	1501	GGCTGATCAAAAAGCGGGCAGCAACTGCAACTGCAACGGCAACACACCTATACGGCC	1560
Db	1444	GGCTGATCAAAAAGCGGGCAGCAACTGCAACTGCAACGGCAACACACCTATACGGCC	1503
Qy	1561	AAACCAATTATCGAAGCGGTTTCGCTGGTGTGTACGCAACAACAATCGGATATGCGC	1620
Db	1504	AAACCAATTATCGAAGCGGTTTCGCTGGTGTGTACGCAACAACAATCGGATATGCGC	1563
Qy	1621	GTGAAACCAAAGTGCCTGATTTATAACGGGGCGGCATCCGCGGTAGCTGGAACAGC	1680
Db	1564	GTGAAACCAAAGTGCCTGATTTATAACGGGGCGGCATCCGCGGTAGCTGGAACAGC	1623
Qy	1681	GACGGCATTTCTATCTGCGAGATACCGNACGATCCGGCGCAACAGGAACGGTCACATC	1740
Db	1624	GACGGCATTTCTATCTGCGAGATACCGNACCAATCCGGCGCAACAGGAACGGTCACATC	1683
Qy	1741	AAAGCGCATCTGCAGCTGGCGGCAAGGTACGCTGTACACAGTTTGGGCAAACTCGTG	1800
Db	1684	AAAGCGCATCTGCAGCTGGCGGCAAGGTACGCTGTACACAGTTTGGGCAAACTCGTG	1743
Qy	1801	AAAGTGGACGGTACGCGGATGACCGCGGCAAGCTGTACATGTTCGGCACCGGCAAGGG	1860
Db	1744	AAAGTGGACGGTACGCGGATTAACGGCGCAAGCTGTACATGTTCGGCACCGGCAAGGG	1803
Qy	1861	GCAGGCTATCTAACCGTACCGGACACGTGTTCCTTTCCTGAGTGCGCCCAAAATCGGG	1920
Db	1804	GCAGGCTATCTAACAGTACCGGACAGCGTTCCTTTCCTGAGTGCGCCCAAAATCGGG	1863
Qy	1921	CGGGATTATTCTTTCTTCAAAACATCGAAACCGACGCTGCTGTCTGGCTTCCTCTCGAC	1980
Db	1864	CAGGATTATTCTTTCTTCAAAACATCGAAACCGACGCTGCTGTCTGGCTTCCTCTCGAC	1923
Qy	1981	AGCGTCGAAAAACACGGGCGAGTGAAGCGCACACGCTGTCTCTATTATGTCCGTCGCGCC	2040
Db	1924	AGCGTCGAAAAACACGGGCGAGTGAAGCGCACACGCTGTCTCTATTATGTCCGTCGCGCC	1983
Qy	2041	AATCGCGACGGACTGTTTCGGCAGGGCACATTCGGCGCCGCGCGGTCTCTGAAACAGCC	2100
Db	1984	AATCGCGACGGACTGTTTCGGCAGGGCACATTCGGCGCCGCGCGGTCTCTGAAACAGCC	2043
Qy	2101	GTAGAACAGGGCGGCAGCAATCTGAAAAACCTGATGTGTGCAACTGGATGCTCTCGGAATCA	2160
Db	2044	GTAGAACAGGGCGGCAGCAATCTGAAAAACCTGATGTGTGCAACTGGATGCTCTCGGAATCA	2103
Qy	2161	TCGCAACACCCGAGACGGTTGAACCTGCGGCCCGCGCACCGACAGATATGCGCGGGCATC	2220
Db	2104	TCGCAACACCCGAGACGGTTGAACCTGCGGCCCGCGCACCGACAGATATGCGCGGGCATC	2163

Qy	2221	CGCCCTACGGCCACATTTTCGCGCAGCGCGCAGCCGTACAGCATGCGAATGCCGCCGAC	2289
Db	2164	CGCCCTACGGCCACATTTTCGCGCAGCGCGCAGCCGTACAGCATGCGAATGCCGCCGAC	2223
Qy	2281	GGTGACGCATCTTCAACAGTCTCGCGCGTACCGTCTATGCCGACAGTACCGCGCCCAT	2340
Db	2224	GGTGACGCATCTTCAACAGTCTCGCGCGTACCGTCTATGCCGACAGTACCGCGCCCAT	2283
Qy	2341	GCGGATATGCAGGAGCCCGCGCTGAAAGCCGTATCGGACGGTTTGGACACCAACGCTACG	2400
Db	2284	GCGGATATGCAGGAGCCCGCGCTGAAAGCCGTATCGGACGGTTTGGACACCAACGCGACG	2343
Qy	2401	GGTCGCGGCTCATCGGCGAACCACAGGACGGTGGAAAGTGGGAACAGGGCGGTGTT	2460
Db	2344	GGTCGCGGCTCATCGGCGAACCACAGGACGGTGGAAAGTGGGAACAGGGCGGTGTT	2403
Qy	2461	GAAGGCAAAATGCGGGCAGTACCCAAACCGCTCGGCATTTCGCGGAAAAACCGCGGAAAT	2520
Db	2404	GAAGGCAAAATGCGGGCAGTACCCAAACCGCTCGGCATTTCGCGGAAAAACCGCGGCAAT	2463
Qy	2521	ACGACAGCCGCCACACTGGGCACTGGGACAGCACAGCACATGGAGCGAAAAACAGTGCAAT	2580
Db	2464	ACGACAGCCGCCACACTGGGCACTGGGACAGCACAGCACATGGAGCGAAAAACAGTGCAAT	2523
Qy	2581	GCAAAAACCGACAGCATTAGTCTGTTTCAGGACATACGGCAGCATCGGCGGATATCGGC	2640
Db	2524	GCAAAAACCGACAGCATTAGTCTGTTTCAGGACATACGGCAGCATCGGCGGCGATATCGGC	2583
Qy	2641	TATCTAAAGCGCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCGCAGCACCCGT	2700
Db	2584	TATCTAAAGCGCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCGCAGCACCCGT	2643
Qy	2701	GCGAGCAACATCGGGAAGCGAGCTCAACGGCAGCTGATGCAGCTGGGCGCACTGGGC	2760
Db	2644	GCGAGCAACATCGGGAAGCGAGCTCAACGGCAGCTGATGCAGCTGGGCGCACTGGGC	2703
Qy	2761	GGTGTAACGCTCCGTTTGCCGCAACGGGAGATTTGACGGTCTGAAGGGGGTCTCGCGTAC	2820
Db	2704	GGTGTAACGCTCCGTTTGCCGCAACGGGAGATTTGACGGTCTGAAGGGGGTCTCGCGTAC	2763
Qy	2821	GACCTGCTCAACAGGATGCATTCGCCGAAAAAGCAGTGTCTTTGGCTGGAGCGGCAAC	2880
Db	2764	GACCTGCTCAACAGGATGCATTCGCCGAAAAAGCAGTGTCTTTGGCTGGAGCGGCAAC	2823
Qy	2881	AGCCTACTCAAGGCACACTGTCGGACATCGCGGTCTGGAAGCTGTGCAACACCTTGAGC	2940
Db	2824	AGCCTACTCAAGGCACACTGTCGGACATCGCGGTCTGGAAGCTGTGCAACACCTTGAGC	2883
Qy	2941	GATAAAGCCGTCCTGTTTGCAACGGGGCGCTGGAACGCGACCTGAACGAGCGGCACTAC	3000
Db	2884	GATAAAGCCGTCCTGTTTGCAACGGGGCGCTGGAACGCGACCTGAACGAGCGGCACTAC	2943
Qy	3001	ACGCTAAACGGGCGGCTTTACCGCGCGACTGCAACCAACCGCAAGACGGGGGCAACGCAAT	3060
Db	2944	ACGCTAAACGGGCGGCTTTACCGCGCGACTGCAACCAACCGCAAGACGGGGGCAACGCAAT	3003
Qy	3061	ATGCCGCACACCCGCTGGTTGCCGGTCTGGGCGGGGATGTCGAATTCGGCAACGGCTGG	3120
Db	3004	ATGCCGCACACCCGCTGGTTGCCGGTCTGGGCGGGGATGTCGAATTCGGCAACGGCTGG	3063
Qy	3121	AACGGCTTGGCAGGTTACAGCTACGCGGGTCTCCAAACAGTACGGCAACACACAGGGGACGA	3180
Db	3064	AACGGCTTGGCAGGTTACAGCTACGCGGGTCTCCAAACAGTACGGCAACACACAGGGGACGA	3123
Qy	3181	GTCGGCGTAGGCTACCGGTTCC 3201	
Db	3124	GTCGGCGTAGGCTACCGGTTCC 3144	

RESULT 18	AX239721	LOCUS	DEFINITION
-----------	----------	-------	------------

linear PAT 26-SEP-2001

ACCESSION AX239721 GI:15797372  
VERSION AX239721.1  
KEYWORDS synthetic construct.  
SOURCE synthetic construct.  
ORGANISM artificial sequences.  
REFERENCE 1 (bases 1 to 4179)  
AUTHORS Arico,M.B., Comanducci,M., Galeotti,C., Masignani,V., Giuliani,M.M.  
and Pizzi,M.  
TITLE Hybrid expression of neisserial proteins  
JOURNAL Patent: WO 0164920-A 23 07-SEP-2001;  
Chiron Spa (JP)  
FEATURES  
    source location/Qualifiers  
        1..4179  
            /organism="synthetic construct"  
            /db\_xref="taxon:32630"  
            /note="deltaG983-961c"  
BASE COUNT 1187 a 1140 c 1121 g 731 t  
ORIGIN  
    Query Match 92.8%; Score 2971.8; DB 6; Length 4179;  
    Best Local Similarity 97.2%; Pred. No. 0;  
    Matches 3052; Conservative 0; Mismatches 47; Indels 42; Gaps 1;  
QY 103 ACTTCTGCCCGGACGTTCAATGACGCGGCACCGGTATCGGCAGCAACAGCAGAGCAACA 162  
DB 4 ACTTCTGCCCGGACGTTCAATGACGCGGCATCGCGGTATCGGCAGCAACAGCAGAGCAACA 63  
QY 163 ACAGCGAATATCAGCAGCAGTATCTTACGCGCGGTATCAAGAACGAAATGTGCAAGACAGA 222  
DB 64 ACAGCGAATATCAGCAGCAGTATCTTACGCGCGGTATCAAGAACGAAATGTGCAAGACAGA 123  
QY 223 AGCATGCTCTGTGCCGCGTCCGGATGACGTTTGGGTTTACAGACAGGGATGCCAAATCAAT 282  
DB 124 AGCATGCTCTGTGCCGCGTCCGGATGACGTTTGGGTTTACAGACAGGGATGCCAAATCAAT 183  
QY 283 GCCCGCCCCCGAATCTGCATACCGGAGACTTTACAAACCCCAAAATGACGCATACAAGAAT 342  
DB 184 GCCCGCCCCCGAATCTGCATACCGGAGACTTTTCCAAACCCCAAAATGACGCATACAAGAAT 243  
QY 343 TTGATCAACTCAAACTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGCT 402  
DB 244 TTGATCAACTCAAACTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGCT 303  
QY 403 ATCTCGTATACAGGCGAATCCGTCGGCAGCATATCCTTTCCGGAATGTATGCGCAAAA 462  
DB 304 ATCTCGTATACAGGCGAATCCGTCGGCAGCATATCCTTTCCGGAATGTATGCGCAAAA 363  
QY 463 GAACAGGCTATACGAAATTTACAAAACCTATACGGCGTATATGCGGAAGGAGCGCCT 522  
DB 364 GAACAGGCTATACGAAATTTACAAAACCTATACGGCGTATATGCGGAAGGAGCGCCT 423  
QY 523 GAAGAGGAGCGGTAAACACATTAAGCTCTTTTCGAGCATGAGCCGCTTATAGAGACT 582  
DB 424 GAAGAGGAGCGGTAAACACATTAAGCTCTTTTCGAGCATGAGCCGCTTATAGAGACT 483  
QY 583 GAAGCAACCGCAGCGATATCGCCACGTTAAAGAAATCGGACACATCGATGTGGTCTCC 642  
DB 484 GAAGCAACCGCAGCGATATCGCCACGTTAAAGAAATCGGACACATCGATGTGGTCTCC 543  
QY 643 CATATATTGGCGGGGTTCCGTGGACGAGACCTGCGAGCGGTATTCGGCCCGATCGG 702  
DB 544 CATATATTGGCGGGGTTCCGTGGACGAGACCTGCGAGCGGTATTCGGCCCGATCGG 603  
QY 703 AGCTACACATTAATGAATACGATGATGAACCAAGAAATATATGTCTGCAGCCATC 762  
DB 604 AGCTACACATTAATGAATACGATGATGAACCAAGAAATATATGTCTGCAGCCATC 663  
QY 763 CGCAATGATGGGTCAAGCTGGCGAAACGTGGCGTCCGCATCGTCAATTAACAGTTTGA 822  
DB 664 CGCAATGATGGGTCAAGCTGGCGAAACGTGGCGTCCGCATCGTCAATTAACAGTTTGA 723  
QY 823 ACAACATCGAGGCGAGGCACTCCGACCATTTTCCAAATAGCCAAATTCGGAGGAGCAGTAC 882

DB 724 ACAACATCGAGGCGAGGCACTGCCGACCTTTTCCAAATAGCCAATTCGGAGGAGCAGTAC 783  
QY 883 CGCCAAGCGTTGCTCGCCTATTCCGGCGGTGATAAACACAGACGAGGTATCCGCCCTGATG 942  
DB 784 CGCCAAGCGTTGCTCGCCTATTCCGGCGGTGATAAACACAGACGAGGTATCCGCCCTGATG 843  
QY 943 CAACACAGCGATTACCGGCAACTTGTCTCCTACCATCCGTAATAAATAACATGCTTTTCATT 1002  
DB 844 CAACACAGCGATTACCGGCAACTTGTCTCCTACCATCCGTAATAAATAACATGCTTTTCATT 903  
QY 1003 TTTTTCGGAAGCAATGACGCACAAGCTCAGCCCAACACACTTGACCTATTGTCATTTTAT 1062  
DB 904 TTTTTCGACAGGCAATGACGCACAAGCTCAGCCCAACACATATGCCCTATTGTCATTTTAT 963  
QY 1063 GAAAGATGCTCAAAAAGGCATTATACAGTCCGAGCGGTAGACCGCAGTGAGAAAAG 1122  
DB 964 GAAAGATGCTCAAAAAGGCATTATACAGTCCGAGCGGTAGACCGCAGTGAGAAAAG 1023  
QY 1123 TTCA-----ATGGCTCCAACCAT 1140  
DB 1024 TTCAACGGGAAATGTATGGAGAACCGGGTACAGAACCCGCTTTCAGTATGGCTCCAACCAT 1083  
QY 1141 TCGGGAATTACTGCCATGTGGTGCCTATCGGCACCCCTATGAAGCAAGCGTCCGTTTCACC 1200  
DB 1084 TCGGGAATTACTGCCATGTGGTGCCTATCGGCACCCCTATGAAGCAAGCGTCCGTTTCACC 1143  
QY 1201 CGTACAAACCCGATTCAAAATGCGCGGAACATCTTTTCGCACCCCATCTGAACCGGCAG 1260  
DB 1144 CGTACAAACCCGATTCAAAATGCGCGGAACATCTTTTCGCACCCCATCTGAACCGGCAG 1203  
QY 1261 GCGGCTCTGCTGTCAGAAATACCCGTGGATGAGCAACGACAACTGCGTACACGCTG 1320  
DB 1204 GCGGCTCTGCTGTCAGAAATACCCGTGGATGAGCAACGACAACTGCGTACACGCTG 1263  
QY 1321 CTGACAAACCGCTCAGACATFCGCTGAGTGGCGTGGACAGCAAGTTCGGCTGGGACTG 1380  
DB 1264 CTGACGACGCTCAGGACATCGGTGACGTGGCGTGGACAGCAAGTTCGGCTGGGACTG 1323  
QY 1381 CTGGATTCGGGTAAAGCCATGAACGACCGCGCTCTTCCGTTTCGGGACTTACCGCC 1440  
DB 1324 CTGGATTCGGGTAAAGCCATGAACGACCGCGCTCTTCCGTTTCGGGACTTACCGCC 1383  
QY 1441 GATACAAAGGTACATCCGATATTCCTTCTCCGTAACGACATTTTCAGGACGCGGC 1500  
DB 1384 GATACAAAGGTACATCCGATATTCCTTCTCCGTAACGACATTTTCAGGACGCGGC 1443  
QY 1501 GGCCTGATCAAAAAGGCGGCGACCAACTGCAACTGCAGGCAACACACCTATATCGGGC 1560  
DB 1444 GGCCTGATCAAAAAGGCGGCGACCAACTGCAACTGCAGGCAACACACCTATATCGGGC 1503  
QY 1561 AAAACCATTAACGAGCGGTTCGCTGGTGTTCGAGGCAACAAATAATCGGATATCGGC 1620  
DB 1504 AAAACCATTAACGAGCGGTTCGCTGGTGTTCGAGGCAACAAATAATCGGATATCGGC 1563  
QY 1621 GTCGAAACCAAGGTTCGCTGATTATTAACGCGGCGGCATCCCGGTAGCCTGAACAGC 1680  
DB 1564 GTCGAAACCAAGGTTCGCTGATTATTAACGCGGCGGCATCCCGGTAGCCTGAACAGC 1623  
QY 1681 GACGCAATTCATCTGCGAGATACCGACCGGATCCGGGCGAAACGAAACCGGTGCACATC 1740  
DB 1624 GACGCAATTCATCTGCGAGATACCGACCAATCCGGGCGAAACGAAACCGGTGCACATC 1683  
QY 1741 AAAGGCGATTCGAGTGGCGCGGAGGTACGCTGTACACAGTTCGGGCAACCTGCTG 1800  
DB 1684 AAAGGCGATTCGAGTGGCGCGGAGGTACGCTGTACACAGTTCGGGCAACCTGCTG 1743  
QY 1801 AAAGTGGAGGTACGCGGATGACCGCGGCAAGCTGTACATGTCGCGACGCGCAAGGG 1860  
DB 1744 AAAGTGGAGGTACGCGGATGACCGCGGCAAGCTGTACATGTCGCGACGCGCAAGGG 1803  
QY 1861 GCAGGCTATCTCAACCGTACCGGACAACTGTTCCCTTCTGAGTCCCGCCCAAAATCGG 1920



Db 1804 GCAGGCTATCTCAACAGTACCGAGACAGTGTTCCTTCTCTAGTGCCGCCCAAAATCGGG 1863  
Qy 1921 CGGGATATTTCTTTCTTCAAAACATCGAAACCGAGCGTGTCTGCTGGCTTCCCTCGAC 1980  
Db 1864 CAGGATATTTCTTTCTTCAAAACATCGAAACCGAGCGGCGCTGCTGGCTTCCCTCGAC 1923  
Qy 1981 AGCGTCGAAAAACAGCGGCGAGTGAAGCGGACAGCTGTCTTATATATCTCGTCCGCGC 2040  
Db 1924 AGCGTCGAAAAACAGCGGCGAGTGAAGCGGACAGCTGTCTTATATATCTCGTCCGCGC 1983  
Qy 2041 AATCGGCGAGGAGTGTTCGCGAGCGGACATTCGCGCGCGCGCGCTGCTGAACACGCG 2100  
Db 1984 AATCGGCGAGGAGTGTTCGCGAGCGGACATTCGCGCGCGCGCGCTGCTGAACACGCG 2043  
Qy 2101 GTAGAACAGGCGGCGAGCAATCTGAAAAACCTGTAGTGTGCAACTGATGCTCCCAATCA 2160  
Db 2044 GTAGAACAGGCGGCGAGCAATCTGAAAAACCTGTAGTGTGCAACTGATGCTCCCAATCA 2103  
Qy 2161 TCCGCAACCCGAGACGGTTTGAAACTCGCGCGCGGCGGACAGATATGCGGGCATC 2220  
Db 2104 TCCGCAACCCGAGACGGTTTGAAACTCGCGCGCGGCGGACAGATATGCGGGCATC 2163  
Qy 2221 CGCCCTTACGGCGCAACTTTCGCGCGAGCGGCGGCTTACAGCATGCAATGCGCGCGAC 2280  
Db 2164 CGCCCTTACGGCGCAACTTTCGCGCGAGCGGCGGCTTACAGCATGCAATGCGCGCGAC 2223  
Qy 2281 GGTGTACGATCTTCAACAGTCTCGCGCTACCGTCTATGCGGACAGTACGCGGCCCAT 2340  
Db 2224 GGTGTACGATCTTCAACAGTCTCGCGCTACCGTCTATGCGGACAGTACGCGGCCCAT 2283  
Qy 2341 GCCGATATGCAAGGACGCGGCTTGAAGCCGCTATCGGACGGTGTGACCAACAGCTACG 2400  
Db 2284 GCCGATATGCAAGGACGCGGCTTGAAGCCGCTATCGGACGGTGTGACCAACAGCGCAG 2343  
Qy 2401 GGTCTGCGGTCATCGCGAAACCAACAGGACGCTGGAACGTGGGAAACAGGCGGCTGT 2460  
Db 2344 GGTCTGCGGTCATCGCGAAACCAACAGGACGCTGGAACGTGGGAAACAGGCGGCTGT 2403  
Qy 2461 GAAGCAAAATGCGGGGAGTACCAACCGTCGCGATTGCGCGCAAAACGCGGCAAAAT 2520  
Db 2404 GAAGCAAAATGCGGGGAGTACCAACCGTCGCGATTGCGCGCAAAACGCGGCAAAAT 2463  
Qy 2521 AGCAGCAGCGCGGCACACTGGGCTATGGGACACAGCATGAGGAGGAAACAGTGCAT 2580  
Db 2464 AGCAGCAGCGCGGCACACTGGGCTATGGGACACAGCATGAGGAGGAAACAGTGCAT 2523  
Qy 2581 CCAAAACCGACAGCATTTAGTCTGTTTCGAGCATACGCGCAGCATGCGGGGATATCGGC 2640  
Db 2524 CCAAAACCGACAGCATTTAGTCTGTTTCGAGCATACGCGCAGCATGCGGGGATATCGGC 2583  
Qy 2641 TATCTCAAGGCTCTTCTCTACGAGCGTACAAAAACAGCATCAGCGGAGCAGCGGT 2700  
Db 2584 TATCTCAAGGCTCTTCTCTACGAGCGTACAAAAACAGCATCAGCGGAGCAGCGGT 2643  
Qy 2701 GCGGACGAACATGCGGAAGGACGCTCAACGCGACGCTGATGAGCTGGGCGCATGGGC 2760  
Db 2644 GCGGACGAACATGCGGAAGGACGCTCAACGCGACGCTGATGAGCTGGGCGCATGGGC 2703  
Qy 2761 GGTGTCAACGTTTCTTTCGCGCAACGCGGAGATTTGACGCTGGAAGGCGGTCTGGCTAC 2820  
Db 2704 GGTGTCAACGTTTCTTTCGCGCAACGCGGAGATTTGACGCTGGAAGGCGGTCTGGCTAC 2763  
Qy 2821 GACCTGCTCAACAGGATGCTTTCGCGCAAAAGGACGCTGCTTGGCTGGAGCGGCAAC 2880  
Db 2764 GACCTGCTCAACAGGATGCTTTCGCGCAAAAGGACGCTGCTTGGCTGGAGCGGCAAC 2823  
Qy 2881 AGCCTACTGAAGGCACACTGGTTCGAGCTCGCGGCTGTAAGCTGTGCAACCCCTTGAGC 2940  
Db 2824 AGCCTACTGAAGGCACACTGGTTCGAGCTCGCGGCTGTAAGCTGTGCAACCCCTTGAGC 2883  
Qy 2941 GATAAAGCGCTCTTTCGCAACGCGGCGGTGGAAACGCGACCTCAACGCGGCGCACTAC 3000  
Db 2884 GATAAAGCGCTCTTTCGCAACGCGGCGGTGGAAACGCGACCTCAACGCGGCGCACTAC 2943

Qy 3001 ACGGTAACGGCGGCTTTACGGCGCGACTCGAGCAACCGGCAACGCGGGGACGCAAT 3060  
Db 2944 ACGGTAACGGCGGCTTTACGGCGCGACTCGAGCAACCGGCAACGCGGGGACGCAAT 3003  
Qy 3061 ATGCGGCACACCGCGCTGTTGCCGCTGCGGCGGATGTGGAATTCGGCAACGGCTGG 3120  
Db 3004 ATGCGGCACACCGCGCTGTTGCCGCGCTGCGGCGGATGTGGAATTCGGCAACGGCTGG 3063  
Qy 3121 AACGCTTGGCACGCTTACAGCTACGCGGTTCCAAACAGTACGGCAACACAGCGGAGCA 3180  
Db 3064 AACGCTTGGCACGCTTACAGCTACGCGGTTCCAAACAGTACGGCAACACAGCGGAGCA 3123  
Qy 3181 GTCGCGTAGGCTACCGGTTT 3201  
Db 3124 GTCGCGTAGGCTACCGGTTT 3144

RESULT 19  
AX236411  
LOCUS AX236411 4344 bp DNA linear PAT 26-SEP-2001  
DEFINITION Sequence 104 from Patent WO0164922.  
ACCESSION AX236411  
VERSION AX236411.1 GI:15796028  
KEYWORDS  
SOURCE synthetic construct.  
ORGANISM synthetic construct.  
REFERENCE 1 (bases 1 to 4344)  
AUTHORS Arico,M.B., Comanducci,M.C., Galeotti,C., Masignani,V.C.,  
Gulliani,M.M. and Pizzi,M.C.  
TITLE Heterologous expression of neisserial proteins  
JOURNAL Patent: WO 0164922-A 104 07-SEP-2001;  
Chiron Spa (IT)

FEATURES  
source  
1. 4344  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="delta983-961"

BASE COUNT 1219 a 1191 c 1167 g 767 t  
ORIGIN

Query Match 92.8%; Score 2971.8; DB 6; Length 4344;  
Best Local Similarity 97.2%; Pred. No. 0;  
Matches 3052; Conservative 0; Mismatches 47; Indels 42; Gaps 1;

Qy 103 ACTTCTGCGCGGCTTCAATGCGGCGGCGGCTATCGGAGCAACAGCAGCAACA 162  
Db 4 ACTTCTGCGCGGCTTCAATGCGGCGGCTATCGGAGCAACAGCAGCAACA 63  
Qy 163 ACAGGAAATCAGCAGCAGTATCTTACCGCGGTATCAAGAACGAAATGTGCAAGACAGA 222  
Db 64 ACAGGAAATCAGCAGCAGTATCTTACCGCGGTATCAAGAACGAAATGTGCAAGACAGA 123  
Qy 223 AGCATGCTCTGCGGCTGCGGATGACGCTTACGAGTACAGACAGGATGCCAAATCAAT 282  
Db 124 AGCATGCTCTGCGGCTGCGGATGACGCTTACGAGTACAGACAGGATGCCAAATCAAT 183  
Qy 283 GCCCGCCCCCGAATCTCGATACCGGAGACTTTCAAAACCCAAATGAGCATACAGAAAT 342  
Db 184 GCCCGCCCCCGAATCTCGATACCGGAGACTTTCCAAACCCAAATGAGCATACAGAAAT 243  
Qy 343 TTGATCAACCTCAACCTGCAATTCGAACGAGCTATACAGCAGCGGCTAGAGTAGGT 402  
Db 244 TTGATCAACCTCAACCTGCAATTCGAACGAGCTATACAGCAGCGGCTAGAGTAGGT 303  
Qy 403 ATCGTCGATACAGGCGAATCGTTCGCGAGCATATCTTTCCGGAATCTGTATGGCAGAAA 462  
Db 304 ATCGTCGATACAGGCGAATCGTTCGCGAGCATATCTTTCCGGAATCTGTATGGCAGAAA 363  
Qy 463 GAACAGGCTATACGAGAAATTTACAAAACCTATACGCGGTATATCGGAGGAAGCGGCT 522  
Db 364 GAACAGGCTATACGAGAAATTTACAAAACCTATACGCGGTATATCGGAGGAAGCGGCT 423







Db	1144	CGTACAAACCCGATTCAAATTCGCCGAACATCCTTTTCGGCACCCATCGTAAACGGCAGC	1203
Qy	1261	GGGCTCTGCTGCTGCAGAAATACCGTGGATGAGCAACAGCAACCTCGGTACCACGCTG	1320
Db	1204	CGGCTCTGCTGCTGCAGAAATACCGTGGATGAGCAACAGCAACCTCGGTACCACGCTG	1263
Qy	1321	CTGACAAACGGCTCAGACATTCGGTGCAGTCGGCTGGGTGGACAGAGTTTCGGCTGGGACTG	1380
Db	1264	CTGACACAGGCTCAGACATTCGGTGCAGTCGGCTGGGTGGACAGAGTTTCGGCTGGGACTG	1323
Qy	1381	CTGGATCGGGTAAAGCCATGAACGACCGCGCTCTTTTCGTTTCGGCGACTTTACCGCC	1440
Db	1324	CTGGATCGGGTAAAGCCATGAACGACCGCGCTCTTTTCGTTTCGGCGACTTTACCGCC	1383
Qy	1441	GATACAAAGGTACATCCGATATGCTACTCTTTCGCTAAACAGCATTTTCAGGCACGGGC	1500
Db	1384	GATACAAAGGTACATCCGATATGCTACTCTTTCGCTAAACAGCATTTTCAGGCACGGGC	1443
Qy	1501	GGCTGATCAAAAAGGCGGAGCAACTGCAACTGCAAGGGAACAACACCTATACGGGC	1560
Db	1444	GGCTGATCAAAAAGGCGGAGCAACTGCAACTGCAAGGGAACAACACCTATACGGGC	1503
Qy	1561	AAACCATATCAAGGCGGTTTCGCTGGTGTGTAAGGCACAACAANAATCGGATATCGC	1620
Db	1504	AAACCATATCAAGGCGGTTTCGCTGGTGTGTAAGGCACAACAANAATCGGATATCGC	1563
Qy	1621	GTGAAACAAAGTCCGCTGATTTATAACGGGCGGCATCCGGCGTAGCTCAACAGC	1680
Db	1564	GTGAAACCAAGTCCGCTGATTTATAACGGGCGGCATCCGGCGTAGCTCAACAGC	1623
Qy	1681	GAGCGATGTCTATCTGGCAGATACCGACCGATCCGGCGGAACAAACCGTCACATC	1740
Db	1624	GAGCGATGTCTATCTGGCAGATACCGACCAATCCGGCGGAACAAACCGTCACATC	1683
Qy	1741	AAAGCGATCTCGAGCTGGCGCGCAAGTACGCTGTACACACGTTTGGGCAAACTGCT	1800
Db	1684	AAAGCGATCTCGAGCTGGCGCGCAAGTACGCTGTACACACGTTTGGGCAAACTGCT	1743
Qy	1801	AAAGTGACGGTACGGCGATGACCGCGCGCAAGCTGTACATGTGGCAGCGCAAAAGG	1860
Db	1744	AAAGTGACGGTACGCGATTTATCGCGGCAAGCTGTACATGTGGCAGCGCGCAAGGG	1803
Qy	1861	GCAGGCTATCTCAACGTTACGGACAACGTGTTCCTTCCTGTAGTGCAGCAAAATCGGG	1920
Db	1804	GCAGGCTATCTCAACAGTACCGGAGAGCGTGTTCCTTCCTGTAGTGCAGCAAAATCGGG	1863
Qy	1921	CGGGATTATTTCTTCTCAACAATCGAAACCGACGCTGTCTGTGGCTTCCTCGAC	1980
Db	1864	CAGGATTATTTCTTCTCAACAATCGAAACCGACGCGGCTGTGTGGCTTCCTCGAC	1923
Qy	1981	AGCGTCGAAAAACACGGCGAGTGAAGCGACACGCTGTCTATTATGTCCGTCCGGCC	2040
Db	1924	AGCGTCGAAAAACACGGCGAGTGAAGCGACACGCTGTCTATTATGTCCGTCCGGCC	1983
Qy	2041	AATGCGGACGGACTGTTTCGGAGGGACAATTCGGCGCCCGCGGTCTGAAACACGCC	2100
Db	1984	AATGCGGACGGACTGTTTCGGAGCGGCACATTCGGCGCCCGCGGTCTGAAACACGCC	2043
Qy	2101	GTAGAACAGGGCGGACGAATCTGGAACACCTGATGTGTGCAACTGGATGCCTCCGAATCA	2160
Db	2044	GTAGAACAGGGCGGACGAATCTGGAACACCTGATGTGTGCAACTGGATGCCTCCGAATCA	2103
Qy	2161	TCCGCAACACCCGAGACGGTTCAAACTCGGGCGCGGACCGCACAGATATGCGGGGATC	2220
Db	2104	TCCGCAACACCCGAGACGGTTCAAACTCGGGCGCGGACCGCACAGATATGCGGGGATC	2163
Qy	2221	CGCCCTTACGGCAACTTTTCGCGCAGCGGACCGGTACAGCATGCGAATGCGCGCGAC	2280
Db	2164	CGCCCTTACGGCAACTTTTCGCGCAGCGGACCGGTACAGCATGCGAATGCGCGCGAC	2223
Qy	2281	GGTGTACGATCTTCAACAGTCTCGCGGCTACCGTCTATGCGCACAGTACCGCGCCAT	2340
Db	2224	GGTGTACGATCTTCAACAGTCTCGCGGCTACCGTCTATGCGCACAGTACCGCGCCAT	2283

Qy	2341	CCCGATATGACGGACGCCGGCTGAAAGCCGCTATCGGACGGGTTGGACCCACAACCGCTACG	2400
Db	2284	GCCGATATGACGGACGCCGGCTGAAAGCCGCTATCGGACGGGTTGGACCCACAACCGCAC	2343
Qy	2401	GGTCTCGCGCTCATCGCGCAAAACCCCAACAGGACGGTGGAACTGGGAACAGGGCGGTGTT	2460
Db	2344	GGTCTCGCGCTCATCGCGCAAAACCCCAACAGGACGGTGGAACTGGGAACAGGGCGGTGTT	2403
Qy	2461	GAAGGCAAAATGCGCGCAGTACCCAAACCGTCGGCATTTGCCGCGAAACACCGCGGAAAT	2520
Db	2404	GAAGGCAAAATGCGCGCAGTACCCAAACCGTCGGCATTTGCCGCGAAACACCGCGGAAAT	2463
Qy	2521	ACGACACGCGCCACACTGGCGATGGGACACACACATGAGCAGCAAAACAGTGCAAT	2580
Db	2464	ACGACACGCGCCACACTGGCGATGGGACACACATGAGCAGCAAAACAGTGCAAT	2523
Qy	2581	GCAAAAACCGACAGCATTAGTCTGTTTGAGGACATACGGCAGCATGCGGGGCGATATCGGC	2640
Db	2524	GCAAAAACCGACAGCATTAGTCTGTTTGAGGACATACGGCAGCATGCGGGGCGATATCGGC	2583
Qy	2641	TATCTCAAAGGCTGTTCTCTACGAGCAGCTACAAAACAGCATCAGCGCAGCACCCGT	2700
Db	2584	TATCTCAAAGGCTGTTCTCTACGAGCAGCTACAAAACAGCATCAGCGCAGCACCCGT	2643
Qy	2701	CGGACCAACATCGGAAGCAGCTCAACGGCACGCTGATCAGCTGGGCGCACTGGGC	2760
Db	2644	CGGACCAACATCGGGAAGCAGCTCAACGGCACGCTGATCAGCTGGGCGCACTGGGC	2703
Qy	2761	GGTGTCAAAGCTGTCGGTTTGGCCCAACGGGAGATTTGACGGTCGAAGCGGTCTGCGCTAC	2820
Db	2704	GGTGTCAAAGCTGTCGGTTTGGCCCAACGGGAGATTTGACGGTCGAAGCGGTCTGCGCTAC	2763
Qy	2821	GACCTGCTCAACAGGATGATTCGCCGAAAAGGCAGTCTTGGCTGGAGCGGCAC	2880
Db	2764	GACCTGCTCAACAGGATGATTCGCCGAAAAGGCAGTCTTGGCTGGAGCGGCAC	2823
Qy	2881	AGCCTCACTGAAGGCACACTGGTCGGACTCGCGGTCTGAAGCTCTCGCAACCCCTTGA	2940
Db	2824	AGCCTCACTGAAGGCACACTGGTCGGACTCGCGGTCTGAAGCTCTCGCAACCCCTTGA	2883
Qy	2941	GATAAGCCCTCTGTTTGACCGCGCGGTGGACGCGACTCAAGCGCGGCACTAC	3000
Db	2884	GATAAGCCCTCTGTTTGACCGCGCGGTGGACGCGACTCAAGCGCGGCACTAC	2943
Qy	3001	ACGCTAACGGCGCGCTTTACCGCGCGACTTGACGAACCGCAACCGGGGCGACGCAAT	3060
Db	2944	ACGCTAACGGCGCGCTTTACCGCGCGACTTGACGAACCGGGGCGACGCAAT	3003
Qy	3061	ATGCCGCACACCCGCTGTTGCCGTCTGGGCGCGGATGTGCAATTCGGCAGCGGCTGG	3120
Db	3004	ATGCCGCACACCCGCTGTTGCCGTCTGGGCGCGGATGTGCAATTCGGCAGCGGCTGG	3063
Qy	3121	AACGGCTTGGCAGTTACAGCTACGCGGTTCCAAACAGTACGGCAACACAGCGGACGA	3180
Db	3064	AACGGCTTGGCAGTTACAGCTACGCGGTTCCAAACAGTACGGCAACACAGCGGACGA	3123
Qy	3181	GTCGGCTAGGCTACCGGTTTC	3201
Db	3124	GTCGGCTAGGCTACCGGTTTC	3144

RESULT 21  
AX236407

LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE

4425 bp DNA linear PAT 26-SEP-2001  
t WO0164922.

AX236407.1 GI:15796026  
synthetic construct.  
synthetic construct  
artificial sequences.  
1 (bases 1 to 4425)

AUTHORS	Arico, M.B., Comanducci, M.C., Galeotti, C.C., Masignani, V.C., Guilian, M.M. and Piza, M.C.									
TITLE	Heterologous expression of neisserial proteins									
JOURNAL	Patent: WO 0164922-A 100 07-SEP-2001;									
	Chiron Spa (IT)									
FEATURES	Location/Qualifiers									
source	1..4425									
	/organism="synthetic construct"									
	/db_xref="taxon:32630"									
	/note="deltaG983-ORF46.1"									
BASE COUNT	1165 a	1259 c	1219 g	782 t						
ORIGIN										
Query Match	92.88;	Score	2971.8;	DB 6;	Length	4425;				
Best Local Similarity	97.2%;	Pred. No.	0;							
Matches 3052;	Conservative	0;	Mismatches	47;	Indels	42;	Gaps			
Qy	103	ACTTCTCGCCCGGCGACTTCAATGATGAGCGCGGACACCGGTATCGCGAGCACACGACGAGCAACA	162							
Db	4	ACTTCTCGCCCGGCGACTTCAATGATGAGCGCGGTACCGGTATCGCGAGCAACACGACGAGCAACA	63							
Qy	163	ACACGGAATACGACGACGATGATCTTACGCCGGTATCAAGAACCAATGTGCAAGACAGA	222							
Db	64	ACACGGAATACGACGACGATGATCTTACGCCGGTATCAAGAACCAATGTGCAAGACAGA	123							
Qy	223	AGCATGCTCTGTGCCGTGGGATGACGTTGTCGGTTACAGACAGGATGCCAAAATCAAT	282							
Db	124	AGCATGCTCTGTGCCGTGGGATGACGTTGTCGGTTACAGACAGGATGCCAAAATCAAT	183							
Qy	283	GCCCCCCCCCGAATCTGCATACCGGACGCTTTACAAACCAATGACGCATACAGAAT	342							
Db	184	GCCCCCCCCCGAATCTGCATACCGGACGCTTTCCAAACCAATGACGCATACAGAAT	243							
Qy	343	TTGATCAACCTCAAACCTGCAATTTGAAGCAGGCTATACAGACGCGGGGTAGAGGTAGGT	402							
Db	244	TTGATCAACCTCAAACCTGCAATTTGAAGCAGGCTATACAGACGCGGGGTAGAGGTAGGT	303							
Qy	403	ATCGTCGATACAGCGGAATCGGTGCGGACGATATCTTTCCCGAATGATGCGCAGAAA	462							
Db	304	ATCGTCGACACAGCGGAATCGGTGCGGACGATATCTTTCCCGAATGATGCGCAGAAA	363							
Qy	463	GAACACGGCTATACGAAAATTCAAAAACTATACGGCTATATCGGAAGGAAGCGCCT	522							
Db	364	GAACACGGCTATACGAAAATTCAAAAACTATACGGCTATATCGGAAGGAAGCGCCT	423							
Qy	523	GAAGACGGAGCGGTAAAGACATTTAAAGTCTCTTTTCGACGATGAGGCCGTTATAGAGACT	582							
Db	424	GAAGACGGAGCGGTAAAGACATTTAAAGTCTCTTTTCGACGATGAGGCCGTTATAGAGACT	483							
Qy	583	GAACAAAGCCGAGGATATCCGCCAGTAAAAGAAATCGGACACATCGATGTGGTCTCC	642							
Db	484	GAACAAAGCCGAGGATATCCGCCAGTAAAAGAAATCGGACACATCGATGTGGTCTCC	543							
Qy	643	CATATTATTGGCGGGCTTCCTGGAGCGGACGCTGCGAGCGGTATTTGGCCCGATGG	702							
Db	544	CATATTATTGGCGGGCTTCCTGGAGCGGACGCTGCGAGCGGTATTTGGCCCGATGG	603							
Qy	703	ACGTACACATAATGAATACGATATGGAACCAAGAACGAAATAATGTCTGCAGCCATC	762							
Db	604	ACGTACACATAATGAATACGATATGGAACCAAGAACGAAATAATGTCTGCAGCCATC	663							
Qy	763	CGCAATCGATGGGTCAAGCTGGGCGACGTTGGCGTCCTCAATACAGTTTTTGA	822							
Db	664	CGCAATCGATGGGTCAAGCTGGGCGACGTTGGCGTCCTCAATACAGTTTTTGA	723							
Qy	823	ACAACTCGAGGGCAGCACTCGCGACATTTCCAAATAGCCAAATTCGGAGGACGATC	882							
Db	724	ACAACTCGAGGGCAGCACTCGCGACATTTCCAAATAGCCAAATTCGGAGGACGATC	783							
Qy	883	CGCCAAAGCTTGTCCGCTATTTCGGGGGTGATTAACACAGCAGGGGTATCCGCTTGATG	942							
Db	784	CGCCAAAGCTTGTCCGCTATTTCGGGGGTGATTAACACAGCAGGGGTATCCGCTTGATG	843							

Db 1924 AGCGTGGAAAAACAGCGGCAGTGAAGCGACACGCTGTCTATTATGTGCTCCGCGGC 1983  
QY 2041 AATGCGGCACGGACTGCTTCGGCAGGGGCACATTTCGCGCCCGCGCTCTGAAACACGGCC 2100  
Db 1984 AATGCGGCACGGACTGCTTCGGCAGGGGCACATTTCGCGCCCGCGCTCTGAAACACGGCC 2043  
QY 2101 GTAGAACAGGGCGGCAGCAATCTGGAACACCTTGATGGTTCGAACCTGGATGCTCCGAATCA 2160  
Db 2044 GTAGAACAGGGCGGCAGCAATCTGGAACACCTTGATGGTTCGAACCTGGATGCTCCGAATCA 2103  
QY 2161 TCCGCAACACCCAGACGGTGTGAACATGCGGCGCGCGACCGCAACAGATATGCCGGGCATC 2220  
Db 2104 TCCGCAACACCCAGACGGTGTGAACATGCGGCGAGCGACCGCAACAGATATGCCGGGCATC 2163  
QY 2221 CGCCCTTACGGCGCAACTTTCGCGCAGCGGCAGCGGTACAGCATGCGAATGCCGGCAG 2280  
Db 2164 CGCCCTTACGGCGCAACTTTCGCGCAGCGGCAGCGGTACAGCATGCGAATGCCGGCAG 2223  
QY 2281 GGTGTACGCATCTTCAACAGTCTCGCGCTACCGTCTATGCCAGACAGTACCGCGGCCAT 2340  
Db 2224 GGTGTACGCATCTTCAACAGTCTCGCGCTACCGTCTATGCCAGACAGTACCGCGGCCAT 2283  
QY 2341 GCCGATATGACGGACCGCGCTGAAAGCCGTATCGGACGGGTGGACCAACACGCTACG 2400  
Db 2284 GCCGATATGACGGACCGCGCTGAAAGCCGTATCGGACGGGTGGACCAACACGCGCACG 2343  
QY 2401 GGTCTCGCGTCAATCGCGCAACCCCAACAGAGCGGTGGAACGTGGCAACAGGCGGTGT 2460  
Db 2344 GGTCTCGCGTCAATCGCGCAACCCCAACAGAGCGGTGGAACGTGGCAACAGGCGGTGT 2403  
QY 2461 GAAGCAAAATGCGCGCAGTACCCAAACCGTCCGGCATTCGCCGAAACCGCGCAAAAT 2520  
Db 2404 GAAGCAAAATGCGCGCAGTACCCAAACCGTCCGGCATTCGCCGAAACCGCGCAAAAT 2463  
QY 2521 ACAGACAGCCGCCACACTGGGCATGGGACACAGACACATGAGACGCAAAACATGTCAAAAT 2580  
Db 2464 ACAGACAGCCGCCACACTGGGCATGGGACAGCAGCAGCATGTGAGCGCAAAACATGTCAAAAT 2523  
QY 2581 GCAAAACCCACAGCATTAGTCTGTGTTCAGGCGATACGGCAGCATACGGCAGCATATCGGC 2640  
Db 2524 GCAAAACCCACAGCATTAGTCTGTGTTCAGGCGATACGGCAGCATACGGCAGCATATCGGC 2583  
QY 2641 TATCTCAAAAGCGCTGTCTCTCCTACGACGCTACAAAAACAGCATACGCCGACACCGGT 2700  
Db 2584 TATCTCAAAAGCGCTGTCTCTCCTACGACGCTACAAAAACAGCATACGCCGACACCGGT 2643  
QY 2701 GCGGACGAACATCGGAAGCAGCGTCAACGGCAGCGTGTATGACGTGGCGCAGCTGGGC 2760  
Db 2644 GCGGACGAACATCGGAAGCAGCGTCAACGGCAGCGTGTATGACGTGGCGCAGCTGGGC 2703  
QY 2761 GGTGTCAACGTTCCGTTTGCAGCGGAGATTGACGCTCAAGCGCTGTCCGCTAC 2820  
Db 2704 GGTGTCAACGTTCCGTTTGCAGCGGAGATTGACGCTCAAGCGCTGTCCGCTAC 2763  
QY 2821 GACCTGTCAACAGGATGATTCGCGCAAAAGGAGTCTTTGGGCTGGAGCGGCAAC 2880  
Db 2764 GACCTGTCAACAGGATGATTCGCGCAAAAGGAGTCTTTGGGCTGGAGCGGCAAC 2823  
QY 2881 AGCCTCACTGAAGGCACACTGTGAGCTCGCGGTCTGAAAGTGTGCAACCTTTGAGC 2940  
Db 2824 AGCCTCACTGAAGGCACACTGTGAGCTCGCGGTCTGAAAGTGTGCAACCTTTGAGC 2883  
QY 2941 GATAAAGCGCTCTGTTTGAACGGCGGTGGAACGCGACCTGAAACGCGACTAC 3000  
Db 2884 GATAAAGCGCTCTGTTTGAACGGCGGTGGAACGCGACCTGAAACGCGACTAC 2943  
QY 3001 ACGGTAACGGCGGCTTTACCGCGCGACTGACAGCAACCGGCAAGACGGGGCACGCAAT 3060  
Db 2944 ACGGTAACGGCGGCTTTACCGCGCGACTGACAGCAACCGGCAAGACGGGGCACGCAAT 3003  
QY 3061 ATCCCGCACACCCGCTGTGTCGGTCTGGGCGCGGATGTCGAATTCGGCAACGGCTGG 3120

Db 3004 ATCCCGCACACCCGCTGTGTCGGGCGCGGATGTCGAATTCGCAACGGCTGG 3063  
QY 3121 AACGGCTTGGCAGCTTACAGCTACGCGGTTCCAAACAGTACGCAACACACGACGACGA 3180  
Db 3064 AACGGCTTGGCAGCTTACAGCTACGCGGTTCCAAACAGTACGCAACACACGACGACGA 3123  
QY 3181 GTCGGCGTAGGCTACCGGTTTC 3201  
Db 3124 GTCGGCGTAGGCTACCGGTTTC 3144  
RESULT 22  
AX239715  
LOCUS AX239715 4425 bp DNA linear PAT 26-SEP-2001  
DEFINITION Sequence 17 from Patent WO0164920.  
ACCESSION AX239715  
VERSION AX239715.1 GI:15797369  
KEYWORDS  
SOURCE synthetic construct.  
ORGANISM synthetic construct  
REFERENCE 1 (bases 1 to 4425).  
AUTHORS Arico,M.B., Comanducci,M., Galeotti,C., Masignani,V., Giuliani,M.M.  
and Pizza,M.  
TITLE Hybrid expression of neisserial proteins  
JOURNAL Patent: WO 0164920-A 17 07-SEP-2001;  
Chiron Spa (IT)  
FEATURES  
source Location/Qualifiers  
1..4425  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="deltaG983-ORF46.1"  
BASE COUNT 1165 a 1259 c 1219 g 782 t  
ORIGIN  
Query Match 92.8%; Score 2971.8; DB 6; Length 4425;  
Best Local Similarity 97.2%; Pred. No. 0;  
Matches 3052; Conservative 0; Mismatches 47; Indels 42; Gaps 1;  
QY 103 ACTTCTGCGCCGACATTCATGTCAGCGGCGACCGGTATCGGCGACACACAGCAGCA 162  
Db 4 ACTTCTGCGCCGACATTCATGTCAGCGGCGGTACCGGTATCGGCGACACACAGCAGCA 63  
QY 163 ACAGCAAAATCAGCAGCAGTATCTTACGCGGTTCAAGAACGAAATGTCAAGACAGA 222  
Db 64 ACAGCAAAATCAGCAGCAGTATCTTACGCGGTTCAAGAACGAAATGTCAAGACAGA 123  
QY 223 AGCATGCTGTGTCGCGGTTCGGATGACGTTGCGGTTACAGACAGGGATGCCAAATCAAT 282  
Db 124 AGCATGCTGTGTCGCGGTTCGGATGACGTTGCGGTTACAGACAGGGATGCCAAATCAAT 183  
QY 283 GCCCGCCCCCGAATCTGCATACCGGAGACTTTACAAACCAATGACGCATACAGAAT 342  
Db 184 GCCCGCCCCCGAATCTGCATACCGGAGACTTTCCAAACCAATGACGCATACAGAAT 243  
QY 343 TTGATCAACCTCAAACTGCAATTGAAGCAGCTATACAGACGCGGGTAGAGTAGT 402  
Db 244 TTGATCAACCTCAAACTGCAATTGAAGCAGCTATACAGACGCGGGTAGAGTAGT 303  
QY 403 ATCGTGATACAGGCGAATCCGTCGCGAGCATATCCCTTTCCCGAAGTGTATGCGAGAAA 462  
Db 304 ATCGTGATACAGGCGAATCCGTCGCGAGCATATCCCTTTCCCGAAGTGTATGCGAGAAA 363  
QY 463 GAACAGGCTATACGAAATTTACAAAACATACGCGGTATATGCGGAAGAGCGCCT 522  
Db 364 GAACAGGCTATACGAAATTTACAAAACATACGCGGTATATGCGGAAGAGCGCCT 423  
QY 523 GAAGCGGAGCGGTAAAGACATTAAAGCTTCTTCACCATGAGGCGGTATAGAGACT 582  
Db 424 GAAGCGGAGCGGTAAAGACATTAAAGCTTCTTCACCATGAGGCGGTATAGAGACT 483  
QY 583 GAAGCAAAAGCGGAGGATATCCGCCACGTAAGAAATTCGACACATCGATGTGGTCTCC 642

Db 484 GAAGCAAAAGCCGACGGATATCCGCCACGTAAGAAATCGGACACATCGATTGGTCTCC 543  
QY 643 CATATTATTGGGGGGTTCCTGGACGCGACACCTGCGAGGGGTATTGCGCCCGATGG 702  
Db 544 CATATTATTGGGGGGTTCCTGGACGCGACACCTGCGAGCGGTATTGCGCCCGATGG 603  
QY 703 AGCTACACATATGAAATACGATGATGGAACCAAGAACGAATAATGTCGACGCCATC 762  
Db 604 AGCTACACATATGAAATACGATGATGGAACCAAGAACGAATAATGTCGACGCCATC 663  
QY 763 CGCAATGATGGTCAAGCTGGCGGAACCTGGCGTGGCATCGTCAATAACAGTTTGA 822  
Db 664 CGCAATGCAATGGTCAAGCTGGCGGAACCTGGCGTGGCATCGTCAATAACAGTTTGA 723  
QY 823 ACAACATCAGGCGAGGCACTCGCGACCATTTCCAAATAGCCAATTCGGAGGAGTAGTAC 882  
Db 724 ACACATCAGGCGAGGCACTCGCGACCTTTCCAAATAGCCAATTCGGAGGAGTAGTAC 783  
QY 883 CGCCAAAGCGTTGCTGCCCTATTCCGGCGGTGATAAAACAGAGAGGGTATCCGCCCTGATG 942  
Db 784 CGCCAAAGCGTTGCTGCCCTATTCCGGCGGTGATAAAACAGAGAGGGTATCCGCCCTGATG 843  
QY 943 CAACAGAGCGATTACGGCAACTTGCTTACCACATCCGTAATAAAACATGCTTTTCATT 1002  
Db 844 CAACAGAGCGATTACGGCAACCTTGCTTACCACATCCGTAATAAAACATGCTTTTCATC 903  
QY 1003 TTTTTCGGCAAGCAATGACGCACAAGCTCAGCCCAACACACTGACCCTATTGCCATTTTAT 1062  
Db 904 TTTTCGACAGGCATGACGCACAAGCTCAGCCCAACACATATGCCCTATTGCCATTTTAT 963  
QY 1063 GAAAAAGATGCTCAAAAGGCAATTATCAGTCGCGAGCGGTAGACCGCATGAGGAAAAAG 1122  
Db 964 GAAAAAGACGCTCAAAAGGCAATTATCAGTCGCGAGCGGTAGACCGCATGAGGAAAAAG 1023  
QY 1123 TTCA-----ATGCTCCACCAAT 1140  
Db 1024 TTCAACAGGGAATGTATGGAACCGGTACAGAACCGCTTGAGTATGGCTCCACACCT 1083  
QY 1141 TCGGGAATTACTGCCATGTGGTGCTTATCGGCACCCCTATGAAGCAAGCTCGGTTTCACC 1200  
Db 1084 TCGGGAATTACTGCCATGTGGTGCTTATCGGCACCCCTATGAAGCAAGCTCGGTTTCACC 1143  
QY 1201 CGTACAAACCCGATTCAAAATGCGGGAACATCCCTTTTCGCGACCCCATCATACCGGCAG 1260  
Db 1144 CGTACAAACCCGATTCAAAATGCGGGAACATCCCTTTTCGCGACCCCATCATACCGGCAG 1203  
QY 1261 GCGGCTCTGCTGCAGAAATACCGTGSATGAGCAACGACAACCTCGTACCAGCTG 1320  
Db 1204 GCGGCTCTGCTGCAGAAATACCGTGSATGAGCAACGACAACCTCGTACCAGCTG 1263  
QY 1321 CTGACAAACGGCTCAGGACATCGGTGCAGTCGCGGTGAGACAGCAAGTTCGGCTGGGACTG 1380  
Db 1264 CTGACAGCGCTCAGGACATCGGTGCAGTCGCGGTGAGACAGCAAGTTCGGCTGGGACTG 1323  
QY 1381 CTGGATGCGGGTAAAGCCATGAAAGGACCCCGCTCTTTTCGTTGCGGACATTTACCGGC 1440  
Db 1324 CTGGATGCGGGTAAAGCCATGAAAGGACCCCGCTCTTTTCGTTGCGGACATTTACCGGC 1383  
QY 1441 GATACGAAAGGTACATCCGATATGCTACTCTTCCGTAACGACATTTTCAGGCAGGGC 1500  
Db 1384 GATACGAAAGGTACATCCGATATGCTACTCTTCCGTAACGACATTTTCAGGCAGGGC 1443  
QY 1501 GGCCTGATCAAAAAAGCGGCGCAACTGCAACTGCGCGCAACCAACACCTATACGGGC 1560  
Db 1444 GGCCTGATCAAAAAAGCGGCGCAACTGCAACTGCGCGCAACCAACACCTATACGGGC 1503  
QY 1561 AAAACCATTTATCGAAGCGGTTCGCTGGTGTGATAGGCAACAACAAATCGGATATGGGC 1620  
Db 1504 AAAACCATTTATCGAAGCGGTTCGCTGGTGTGATAGGCAACAACAAATCGGATATGGGC 1563  
QY 1621 GTCGAAACCAAGGTGCGCTGATTTATACGGGGCGGATCCGGGTAGCTGACACG 1680  
Db 1564 GTCGAAACCAAGGTGCGCTGATTTATACGGGGCGGATCCGGGTAGCTGACACG 1623

QY 1681 GACGCATTGTCTATCTGSCAGATACCGACCGATCCGGCGCAACAAACCGTGCACATC 1740  
Db 1624 GACGCATTGTCTATCTGSCAGATACCGACCAATCCGGCGCAACAAACCGTGCACATC 1683  
QY 1741 AAAGCGATCTCGAGTGGCGGCGAAGTACGCTGTACACAGTTTGGGCAAACTGCTG 1800  
Db 1684 AAAGCGATCTCGAGTGGCGGCGAAGTACGCTGTACACAGTTTGGGCAAACTGCTG 1743  
QY 1801 AAAGTGGAGGTACCGGATGACCGGCAAGCTGTACATCTCGSCAGCGCGCAAGGG 1860  
Db 1744 AAAGTGGAGGTACCGGATGATCGGCGCAAGCTGTACATCTCGSCAGCGCGCAAGGG 1803  
QY 1861 GCAGGCTATCTCAACCGTACCGGACAACTGTTCCCTTCTTCTGAGTCCGCCAAAATTCGG 1920  
Db 1804 GCAGGCTATCTCAACAGTACCGGACAGCTGTTCCCTTCTTCTGAGTCCGCCAAAATTCGG 1863  
QY 1921 CGGGATATCTTCTTCTCACAAACATCGAAACCGAGCGGTGCTGCTGGCTTCCCTCGAC 1980  
Db 1864 CAGGATATCTTCTTCTCACAAACATCGAAACCGAGCGGCGCTGCTGGCTTCCCTCGAC 1923  
QY 1981 AGCGTCGAAAAACAGCGGCGAGTGAAGCGACACGCTGTCTTATATCTCCGTCCGCG 2040  
Db 1924 AGCGTCGAAAAACAGCGGCGAGTGAAGCGACACGCTGTCTTATATCTCCGTCCGCG 1983  
QY 2041 AATGGCGCAGGACTGCTTTCGGCAGCGGCACATTTCCGCCCGCGCGGTCTGAAACACAGCC 2100  
Db 1984 AATGGCGCAGGACTGCTTTCGGCAGCGGCACATTTCCGCCCGCGCGGTCTGAAACACAGCC 2043  
QY 2101 GTAGAACAGGGCGGAGCAATCTGGAACACCTGTAGTGTGGAACCTGATGCTCCGAATCA 2160  
Db 2044 GTAGAACAGGGCGGCAATCTGGAACACCTGTAGTGTGGAACCTGATGCTCCGAATCA 2103  
QY 2161 TCCGCAACACCGACAGCGTTGAACTCGCGCGCGGACGACAGATATCGCGGCATC 2220  
Db 2104 TCCGCAACACCGACAGCGTTGAACTCGCGCGCGGACGACAGATATCGCGGCATC 2163  
QY 2221 CGCCCTTACGGCGCAACTTTCCGCCAGCGGAGCGGTACAGCATGCGGAATGCCGCGAC 2280  
Db 2164 CGCCCTTACGGCGCAACTTTTCGCCAGCGGAGCGGTACAGCATGCGGAATGCCGCGAC 2223  
QY 2281 GGTGTACGATCTTCAACAGTCTCCCGGTACCGTCTATGCCAGAGTACCGCGCGCAT 2340  
Db 2224 GGTGTACGATCTTCAACAGTCTCCCGGTACCGTCTATGCCAGAGTACCGCGCGCAT 2283  
QY 2341 GCGGATATCGAGGAGCGCGGTGAAAGCGGTATCGAGCGGTGGACCAACAGCTAGC 2400  
Db 2284 GCGGATATCGAGGAGCGCGGTGAAAGCGGTATCGAGCGGTGGACCAACAGCGCAG 2343  
QY 2401 GGTCTCGCGCTCATCGCGCAAAACCAAGAGCGGTGGAACGTGGGAACAGGGCGGTGTT 2460  
Db 2344 GGTCTCGCGCTCATCGCGCAAAACCAAGAGCGGTGGAACGTGGGAACAGGGCGGTGTT 2403  
QY 2461 GAAAGCAAAATCGCGGAGTACCCAAACCGTGGCATTTGCCGGAACAAACCGCGGAAAT 2520  
Db 2404 GAAAGCAAAATCGCGGAGTACCCAAACCGTGGCATTTGCCGGAACAAACCGCGGAAAT 2463  
QY 2521 ACGACAGACGCGCACACTGGGCATGGGACACAGCAGCATGGAGGAAACAGTGCAAAT 2580  
Db 2464 ACGACAGACGCGCACACTGGGCATGGGACAGCAGCATGGAGGAAACAGTGCNAAT 2523  
QY 2581 GCAAAACCGCAGCATTTAGTCTGTTTTCAGGCATACGGCAGCATCGCGGCGATATCGGC 2640  
Db 2524 GCAAAACCGCAGCATTTAGTCTGTTTTCAGGCATACGGCAGCATCGCGGCGATATCGGC 2583  
QY 2641 TATCTCAAAAGCGTGTCTCTTACGAGCGCTACAAAACAGCATCAGCGCAGACCGGT 2700  
Db 2584 TATCTCAAAAGCGTGTCTCTTACGAGCGCTACAAAACAGCATCAGCGCAGACCGGT 2643  
QY 2701 GCGGACGAACTGCGGAGGAGCGGTCAACGCGACGCTGATCAGTGGCGGCGACTGGGC 2760  
Db 2644 GCGGACGAACTGCGGAGGAGCGGTCAACGCGACGCTGATGCTGAGCTGGGCGACTGGGC 2703

QY	2761	GGTGTCAAGCTTCCGTTTCCCGCAACGGGAGATTTCACGGTCCGAAGCGGCTTGCCTAC	2820
Db	2704		
QY	2821	GACCTCTAAACAGATGATTCGCCGAAAGGACAGTCTTTGGGCTGGAGCGGCAAC	2880
Db	2764		
QY	2881	AGCCTCACTGAAGGCACACTGGTTCGGACTCGCGGCTCTGAAGCTGTGCAACCTTTGAGC	2940
Db	2824		
QY	2941	GATAAAGCCGCTCTGTTTGCACAGCGGCGCGTGGAAACGGACCTGAAACGACGCGACTAC	3000
Db	2884		
QY	3001	ACGGTAACGGCGGCTTTACCGCGCGACTGCAGCAACCGGCAAGACGGGGGCACCAAT	3060
Db	2944		
QY	3061	ATGCCGCACACCCGCTGCTGGTTCGGCTCTGGCGCGGATGTGCAATTTCGGCAACGGCTGG	3120
Db	3004		
QY	3121	AACGGCTTGGCAGCTTACAGCTACGCCGCTTCCAAACAGTACGGCAACACACAGCGGACGA	3180
Db	3064		
QY	3181	GTCCGGGTAGGTACCGGTTTC 3201	
Db	3124		
RESULT 23			
AF169448/c			
LOCUS	AF169448	600 bp	DNA linear BCT 09-AUG-2000
DEFINITION	Neisseria meningitidis strain Z2491 clone Cm024 unknown sequence.		
ACCESSION	AF169448		
VERSION	AF169448.1	GI:9754656	
KEYWORDS	Neisseria meningitidis.		
SOURCE	Neisseria meningitidis.		
ORGANISM	Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;		
REFERENCE	1 (bases 1 to 600)		
AUTHORS	Perrin,A., Nassif,X. and Tinsley,C.R.		
TITLE	Identification of regions of the chromosome of Neisseria meningitidis and Neisseria gonorrhoeae which are specific to pathogenic Neisseriae		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 600)		
AUTHORS	Perrin,A., Nassif,X. and Tinsley,C.R.		
TITLE	Direct Submission		
JOURNAL	Submitted (16-JUL-1999) Necker Medicine Faculty, INSERM U411, 156 rue de Vaugirard, Paris 75015, France		
FEATURES	Location/Qualifiers		
source	1. .600		
	/organism="Neisseria meningitidis"		
	/strain="Z2491"		
	/db_xref="taxon:487"		
	/clone="Cm024"		
BASE COUNT	111 a	159 c	144 g 184 t 2 others
ORIGIN			
Query Match	18.0%; Score 576.4; DB 1; Length 600;		
Best Local Similarity	98.5%; Pred. No. 1.1e-121;		
Matches	591; Conservative	0; Mismatches	8; Indels 1; Gaps 1;
QY	306	CGGAGACTTTTACAAACCCCAATCAGCATACAGATTTTGATCAACCTCAACCTGCAAT	365
Db	600		
QY	366	TGAAGCAGGCTATACAGGACGGGGGTAGAGGTAGGTATCGTTCGATACAGGGAATCCGT	425

Db	540	TGAAGCAGGCTATACAGGACGGGGTAGAGTAGTATCGTCGACACAGCGGAATNAGN	481
QY	426	CGGCACATATCCTTTCCGAACTGTATGCGCAAAAGAACACGGCTATAACGAAATTA	485
Db	480		
QY	486	CAAAAACCTATACGGCTATATCGGAAGAACGCCCTGAAGACGGAGCGGTAAAGACAT	545
Db	420		
QY	546	TAAAGCTTCTTTGACGATGAGCGCTTATAGAGACTGAAGCAAAGCCGACGGATATCCG	605
Db	360		
QY	606	CCACGTAAAAGAAATCGGCACATCGATGGTCTCCCATATATTATTTGGGGGGGTTCCGT	665
Db	300	CCACGTAAAAGAAATCGGCACATCGATGGTCTCCCATATATTATTTGGGGGGGTTCCGT	241
QY	666	GGACGCAGACCTGCGAGCGGTATTGCGCCGATGCGAGCTACACATATGAATACGCA	725
Db	240	GGACGCAGACCTGCGAGCGGTATTGCGCCGATGCGAGCTACACATATGAATACGCA	181
QY	726	TGATGGAACCAAGAACAAATATGTCTCAGCCATCCGCAATGCATGGGTCAAGCTGG	785
Db	180	TGATGGAACCAAGAACAAATATGTCTCAGCCATCCGCAATGCATGGGTCAAGCTGG	121
QY	786	CGAACGTGGCGTGGCATCGTCAATACAGTCTTGTGAAACAACATCGAGGCGAGGCACTGC	845
Db	120	CGAACGTGGCGTGGCATCGTCAATACAGTCTTGTGAAACAACATCGAGGCGAGGCACTGC	61
QY	846	CGACCAATTTCCAAATAGCCAATTCGGAGGAGTAGTACCGCAAGCGTTCTCGCCCTATTC	905
Db	60	CGACCAATTTCCAAATAGCCAATTCGGAGGAGTAGTACCG-CAAGCGTTGCTCCCTATCC	2
RESULT 24			
AF169473			
LOCUS	AF169473	286 bp	DNA linear BCT 09-AUG-2000
DEFINITION	Neisseria meningitidis strain Z2491 clone Em085 unknown sequence.		
ACCESSION	AF169473		
VERSION	AF169473.1	GI:9754681	
KEYWORDS	Neisseria meningitidis.		
SOURCE	Neisseria meningitidis.		
ORGANISM	Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;		
REFERENCE	1 (bases 1 to 286)		
AUTHORS	Perrin,A., Nassif,X. and Tinsley,C.R.		
TITLE	Identification of regions of the chromosome of Neisseria meningitidis and Neisseria gonorrhoeae which are specific to pathogenic Neisseriae		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 286)		
AUTHORS	Perrin,A., Nassif,X. and Tinsley,C.R.		
TITLE	Direct Submission		
JOURNAL	Submitted (16-JUL-1999) Necker Medicine Faculty, INSERM U411, 156 rue de Vaugirard, Paris 75015, France		
FEATURES	Location/Qualifiers		
source	1. .286		
	/organism="Neisseria meningitidis"		
	/strain="Z2491"		
	/db_xref="taxon:487"		
	/clone="Em085"		
BASE COUNT	88 a	73 c	63 g 62 t
ORIGIN			
Query Match	8.5%; Score 270.8; DB 1; Length 286;		
Best Local Similarity	99.0%; Pred. No. 2e-51;		
Matches	283; Conservative	0; Mismatches	2; Indels 1; Gaps 1;
QY	865	AATTCGGAGGACGATACCGCAAGCGTTGCTCGCCCTATTTCCGGCGGTATATAACAGAC	924









```
/gene="czcA"
/notice="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer/Genemark"
/codon_start=1
/transl_table=11
/product="cation efflux system protein"
/protein_id="AAM37000.1"
/db_xref="GI:21108375"
/translation="MIEAIIITTCFRRGIVMLVFVLIATLGTWSWTOLPVEAYPDIAAD
VTSOVTVQVGLGAEVEEQITVTERALMGTPGLHVLRSRSLFALSILITLVDFDGHAE
GYFARQVLERIQAVTLPLPYGAIPGLDPTSTPGIYRYTLESKTRSRLEISLDQFWIV
IPRQKVPGVADVTFNGGLTTFQSLALEPDRDLTRIGVSLQOVKSAITSNNADGGSGVM
DRGQSVYVIGLHSLQDIGNVYSSNGVPLVKDLGEVRYDNERGILGKDN
PDTTEGALLKDSNPVALQGHSAVEELNSALPKDKVVPYLDRTALITLALTMHVT
SATITEGMLLVVVLLIFLPGSPRAAAIVSLTIPLSLITAFIVMHLKIIPANLISLGA
DFGLIDGAVVLVENVLRLENSORALTARDAIDATLOVAPRPFEGMAVIGCAVLP
LAFERIEYKLESPNAYAVGAALIGALLVALALIPGLAWLAFPRKRMHNRYRELQ
IRAVLRSVGRGWLACAAALCALCVLALLGSGIRDFLPYIDEGSLWLVQMPPIGT
LDKAATMANALRKATLEFEVSYVVTQGRNDGTDYTSFHSIEASGLRYPKRWPG
MDKALLTAALGARTARMPGYTVSMQVQDKLSAHSDDLTVKVFGLDQOVRGV
ADOVAFAHLKVPGAADTAVDEPPPLPNLQVRFDEAAARYGINAADVSDLISTGIGS
PIGOMYLGOKSYDLTVRPORYRNDPOAIGALRLRTAAGAEIPLSAVANITATTSQSV
IVREGRNIIIVRLNVRGRDLSFSLSDAQATLARQVHLDPQHMQLVWGQFENLQRAQ
ARLVVLPITLICIMFVLFGAFGNLRQPTLVLAAPVPLAMIGGLAALHLRGMTLNVSSA
VGFIALGVAVLNAVLMQAQINRLRQDVGMSLREAVVAGAVSRMRPVLMTATVAALGL
APAMLATGLGSDVQRLATVVVGLVLTATLITLVLLPSLYLMEAFVATRQGRQTFP
".
7877..9361
/gene="ttgC"
/notice="XAC2148"
7877..9361
/gene="ttgC"
/notice="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer/Genemark"
/codon_start=1
/transl_table=11
/product="outer membrane efflux protein"
/protein_id="AAM37001.1"
/db_xref="GI:21108376"
/translation="MTDVRNALASAVAASLAACAVGPDFORPAPAADRYTANALPV
TTAADVHGAQOEFEPCQELPAWWRFGSPOLNAVVERALRNPDLQADALPRA
QOTLAQNGAWRPQADRLVDASRDORSVPVAPDAGIDTPYALRTAQLSVSTLVDFG
TRROVEAQAQAQRDQDAAYLSLTANVVNAIAGEAAVRAQLDAAQVAQVIAKQLC
LTARQAGALGAQAQVLAQRTALAAQAQAAVPALEKLEQOQRTLLAILGGOLPGDAAG
EAFPLDTLQPLPAQLPLSPRLDRPDIRAAEAQAHAASALIGVAAAALPSPISLSA
ASGSGASGVGRAGNTQWSVAGCLLQPLPHFGALAHQORAAEAGVTOAMAQVRSV
LNAFEGVSNALTALQSDAQLHACADAQAELRTALVRQQAVGAVGMADLLQAQA
YQSTQATVQAQARYTDTVALYQALGGPWPSPAARTPKPVP"
9129..10502
/gene="nodV"
/notice="XAC2149"
9129..10502
/gene="nodV"
/notice="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer/Genemark"
/codon_start=1
/transl_table=11
/product="nodulation protein"
/protein_id="AAM37002.1"
/db_xref="GI:21108377"
/translation="MRKRWRSAWRCWDVNRRLVRWAWRCCRSTRPRARRPPRSRH
TSRPTPTPWRIRPWVAGRGQRPARSPRFHESRAASAGCVTAAPPAETIAAMRM
TSLRSPPTTRFDLGMVGLVAGMAAIFVADTFTDYAVAAALFTVLVLLSVRLLTRG
VVLGLRDLVLIVLSFHLTPAGSVRGVINSIVSVVAVTLVAVOMAAKASAHAA
QALLRLTRSGMGLASSIAHEVNOPLAIVTSCNACORWLAQOPLNKAALAAH
MLADAGRASIIARMRALARGEASQRPDLNRTVLDMIALTRGETERHGTALSLELD
PALPVLADGVQIQOVIGNLNLNAVEALATLPDDEARSIQIRKSHIDGAWACLCIDDTGP
GLSDADLHAFVEAFWMTKDDGMGTGLSISRTIMEANGQIWAEPAGVGRFGLRFLP
VATERRO"
10499..11110
/gene="nodW"
/notice="XAC2150"
10499..11110
```

```
/gene="nodW"
/notice="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer/Genemark"
/codon_start=1
Query Match 4.1%; Score 132.4; DB 1; Length 11292;
Best Local Similarity 55.7%; Pred. No. 1.7e-19;
Matches 303; Conservative 0; Mismatches 226; Indels 15; Gaps 2;
QY 1087 ATCACAGTCGCGAGCGTAGACCGCAGTCAGGAGAAAAGTTCAATGGTCCCAACCATTTGCCGA 1146
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1239 ATCAATATCGATAGCGCGGGAATGCCCGCGCTGACGACGATATTCCAATCATTTGCCGC 1298
QY 1147 ATTACTGTCATGTGTGCTATCGGCAC-----CTATGAAGCAAGCGTCCGTTTTCACC 1200
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1299 GCCGCTGCGCAGTGTGCTGCTGCTGCGCGCGGCACGTATACGCCCGCGCTGCGCCGC 1358
QY 1201 CGTACAAACCGGATTTAAATTCGCGGAACATCTCTTTCCGACCCCATCGTAACCGGCAG 1260
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1359 ACCGAATTCGAGGGCGAGATTCGCGGCACCTCGTTCTCCACCGCGCGCTCAGCGGTGTC 1418
QY 1261 GCGGCTCTGCTGCACAAATACCCGTGGATGACGACGACCAACCTGGTACCACGCTG 1320
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1419 GCTGCGCAGGTGCTGGGCGTGTATCCGTGGATGAGCGCTTCCAACCTGCAGCAACCTG 1478
QY 1321 CTGACAACGGCTCAGGACCATCGGTGCAGTCGCGCTGGACAGCAAGTTGCGGTGGGACTG 1380
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1479 CTGACCACCGCACCGACCTGGGCGACCGCGGTGGACGCTGTACGCTGGGCGCTG 1538
QY 1381 CTGATGCGGGTAGGCCATGAACGGACCGCGGTCTTTCGTTCCGTCGCGGACTTTTACC 1440
DB 1539 GTCATATGCGGCGCAAGGCGATCAAGGGCGCGGCCAGTT-----GCGAAGCAACTGG 1589
QY 1441 GATAGGAAGGTACATCCGATATTCCTTACCTTCCCTTAAGACATTCAGGACGCGG 1500
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1590 GCGGCAATGTGACCGCTGGCTACGACAGCACCTTCACGAACGACATCTCCGGCGCGGC 1649
QY 1501 GGCCTGTATCAAAAAAGCGCGGCAACTGCAACTGCGACGCAACAACCATATACGGGC 1560
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1650 AGCCTGACCAAGCGCGCAGGCAAGCTCACCTGTAGCGGCACCAATACCTATATAGCGC 1709
QY 1561 AAACCATTTATCGAAGCGGTTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1620
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1710 GGCACCAAGTTCAGGATGGGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1769
QY 1621 GTCG 1624
DB 1770 GTGG 1773
RESULT 28
AE012308/c
LOCUS
DEFINITION
Xanthomonas campestris pv. campestris str. ATCC 33913, section 216
of 460 of the complete genome.
ACCESSION
AE012308 AE008922
VERSION
AE012308.1 GI:21113147
SOURCE
Xanthomonas campestris pv. campestris str. ATCC 33913.
ORGANISM
Xanthomonas campestris pv. campestris str. ATCC 33913
Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
Xanthomonas.
REFERENCE
1 (bases 1 to 11019)
da Silva,A.C.R., Ferro,J.A., Reinach,F.C., Fatah,C.S., Furlan,L.R.,
Quaggio,R.B., Monteiro-Vitorello,C.B., Van Sluys,M.A., Almeida
Jr.,N.F., Alves,L.M.C., do Amaral,A.M., Bertolini,M.C.,
Camargo,L.E.A., Camarotte,G., Cannavan,F., Cardoso,J.,
Chamberg,F., Clapina,L.P., Cicarelli,R.M.B., Coutinho,L.,
Cursino-Santos,J.R., El-Dorry,H., Faria,J.B., Ferreira,A.J.S.,
Ferreira,R.C.C., Ferro,M.I.T., Formighieri,E.F., Franco,M.C.,
Greggio,C.C., Gruber,A., Katsuyama,A.M., Kishi,L.T., Leite
Jr.,R.P., Lemos,E.G.M., Lemos,M.V.F., Locati,E.C., Machado,M.A.,
Madeira,A.M.B.N., Martinez-Rossi,N.M., Martins,E.C., Meidanis,J.,
```



```
/product="sensor histidine kinase"
/protein_id="AAM41319.1"
/db_xref="GI:21113153"
/translacion="MRSSAPSALNATAIERISAVPOILETVAHITGSRFTAVRVT
DTSWVACITVDITGLFGKLGQLEVESTICHELRSKPVIFFSHASQHPVIAHHTEPK
LYGESVYPIFRDGGFTGTCADISAPANFEADHILKLSLEFALLIGNLSLETEK
LQQAATLAAANDAGELRDQFIAYLGHDLRSPLOSI SMATEMLQLEQSPRGASLLGH
IRTSLTIRGLVDVDFARGRLGGGIPVTLRVEDALERLLEVVQVCSATGRTDLH
ADIHGIEICVCDPRLSOLFANLNAHVCTPSIPVTLRAWQAAGLILHLSVHNGVVI
APEKRALFSPFSDDDDDAPQGLGLGYIAAEAKAHEGRLSVSSOLLAGTTFTFEM
VPAQAQPHQSHVAGDEADALADDARPMOG"
BASE COUNT      1843 a   3451 c   3851 g   1874 t
ORIGIN
```

```
Query Match          4.1%; Score 130.2; DB 1; Length 11019;
Best Local Similarity 57.1%; Pred. No. 5.4e-19;
Matches 284; Conservative 0; Mismatches 198; Indels 15; Gaps 2;

Qy 1132 TCCAAACCATTCGGGAATTACTGCCATGTGGTCCCTATTCGGCACCC-----CTATGAAGCA 1185
Db 1980 TCCAATCATTCGGGTCCGCTGCGCAGTGGTCCCTGTGGCACCGGCGCATGTACGCCGCC 1921

Qy 1186 AGCGTCGCTTTACCCGCTACAAACCCGATTCAAATTGCGGGAACATCCTTTTCCGCACCC 1245
Db 1920 CCGGCCCTGGCGGCACCGAGTTGCAGGCGCCAGATCGCGCGCACCTCGTTCTCCACTGCC 1861

Qy 1246 ATCGTAACCGCGACGCGGCTCTGCTGTCGACGAATACCCGTGGATGAGCAACGACAAAC 1305
Db 1860 GCTGTACGCGGTGGCGCGCAGGACTCTGCGGTGTATCCCTGTGATGAGCGCTTCCCAAC 1801

Qy 1306 CTGCGTACACGCTGCTGACAAACGGCTCAGGACATCGCTGCGGTGGACAGCAAG 1365
Db 1800 CTGCAGACAGCCCTGCTGACCAACGCCACCGACCTGGCGGATCCGCGGTGACGCGGTG 1741

Qy 1366 TTCGCTGGGACTGCTGATGCGGGGTGAAGCCATGAACGACCGCGCTCTTTCCGTTTC 1425
Db 1740 TACGGTTGGGCGATGTTCAATGCCGCCAAGCGATCAAGGGCGCGGCGAGTT----- 1688

Qy 1426 GCGGACTTTACCGCGATACGAAGGTACATCCGATATTGCTACTCTCCGTAAACGAC 1485
Db 1687 --CGCCAGCACTGGCGGGAACCTGTACCTCCGCTACGACAGCACCTTCAGCAACGAC 1630

Qy 1486 ATTTCCAGCAGCGGCGCTGATCAAAAAGGCGGACGCCAACTGCAACTGCACGCGCAAC 1545
Db 1629 ATCTCCGCGAGCGCAACTGTACCAAGAAGCGCGCGCAGACTCACCTTGAGCGGCAAC 1570

Qy 1546 AACACTATACGGGCAAAACCATTTATCGAAGCGGTTCGCTGGTGTGTAGCGGCAACAC 1605
Db 1569 AACAGCTACGCGGCGGACCAACGATCAGCGACGCGTGTGCTGCTTGTTCGGGCGAGCTG 1510

Qy 1606 AAATCGGATATGCGGT 1622
Db 1509 CGCTCCGAGCTGCTGT 1493
```

```
RESULT 29
NGOPAK          NGOPAK          1434 bp   DNA   linear   BCT 10-APR-1995
LOCUS           N.gonorrhoeae opak gene for opacity protein.
DEFINITION      X52364 S36068 S75784
ACCESSION       X52364.1 GI:48702
VERSION         opacity protein; opak gene.
KEYWORDS        Neisseria gonorrhoeae.
SOURCE          Neisseria gonorrhoeae.
ORGANISM        Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
                Neisseria.
REFERENCE       1 (bases 1 to 1434)
AUTHORS        Bhat,K.
TITLE          Direct Submission
JOURNAL        Submitted (31-MAR-1990) Bhat K., National Institutes of Health,
                Rocky Mountain Laboratory, Hamilton 59840, U S A
REFERENCE       2 (bases 1 to 1318)
AUTHORS        Bhat,K.S., Gibbs,C.P., Barrera,O., Morrison,S.G., Jahnig,F.,
```

```
TITLE
JOURNAL          Stern,A., Kupsch,E.M., Meyer,T.F. and Swanson,J.
MEDLINE          The opacity proteins of Neisseria gonorrhoeae strain MS11 are
PUBMED           encoded by a family of 11 complete genes
1815562
192114767
AUTHORS          Mol. Microbiol. 5 (8), 1889-1901 (1991)
REFERENCE        3 (bases 1 to 1434)
                Bhat,K.S., Gibbs,C.P., Barrera,O., Morrison,S.G., Jahnig,F.,
                Stern,A., Kupsch,E.M., Meyer,T.F. and Swanson,J.
TITLE            The opacity proteins of Neisseria gonorrhoeae strain MS11 are
JOURNAL          encoded by a family of 11 complete genes
MEDLINE          Mol. Microbiol. 6 (8), 1073-1076 (1992)
PUBMED           92261323
1584024
COMMENT          The cloned gene is in its non-expressed state. The change in number
                of the coding repeat unit has caused the atg start codon in the
                signal peptide to be out of frame with the rest of its coding
                sequence.
```

```
FEATURES
source          1..1434
                /organism="Neisseria gonorrhoeae"
                /strain="MS11"
                /db_xref="taxon:485"
repeat_unit     218..243
repeat_unit     344..369
-35_signal      373..380
-10_signal      398..403
RBS             455..460
gene            465..1255
                /gene="opak"
                /gene="opak"
CDS             465..1255
                /note="The cloned gene is in its non-expressed state. The
                change in the number of coding repeat unit has caused the
                atg start codon to be out of frame with the rest of its
                coding sequence"
                /pseudo
                /codon_start=1
                /transl_table=11
                /product="opacity protein"
                /db_xref="PID:el4598"
repeat_region   490..524
                /note="signal peptide coding sequence"
                /rpt_type=TANDEM
                /rpt_type=TANDEM
mat_peptide     539..1252
                /gene="opak"
                /product="opacity protein"
terminator      1269..1294
BASE COUNT      383 a   425 c   343 g   283 t
ORIGIN
```

```
Query Match          2.5%; Score 81.6; DB 1; Length 1434;
Best Local Similarity 95.5%; Pred. No. 7e-08;
Matches 84; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3117 CTGGAACGCTTGGCACCTTACAGCTACGCGGTTCACAAAGTACGCAACACGCGG 3176
Db 1 CAGGAACGCTTGGCACCTTACAGCTACGCGGTTCACAAAGTACGCAACACGCGG 60

Qy 3177 ACGAGTCGCGGTAGCTACCGGTTCTGA 3204
Db 61 ACAATCGCGGTAGCTACCGGTTCTGA 88

RESULT 30
AE011766
LOCUS           Xanthomonas axonopodis pv. citri str. 306, section 144 of 469 of
DEFINITION      the complete genome.
ACCESSION       AE011766 AE008923
VERSION         AE011766.1 GI:21107504
KEYWORDS
SOURCE          Xanthomonas axonopodis pv. citri str. 306.
```

## ORGANISM

Xanthomonas axonopodis pv. citri str. 306  
Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;  
Xanthomonas.

REFERENCE  
AUTHORS

1 (bases 1 to 10807)  
da Silva,A.C.R., Ferro,J.A., Reinach,F.C., Farah,C.S., Furlan,L.R.,  
Quaggio,R.B., Monteiro-Vitorello,C.B., Van Sluys,M.A., Almeida  
Jr.,N.F., Alves,L.M.C., do Amaral,A.M., Bertolini,M.C.,  
Camargo,L.E.A., Camarotte,G., Cannavan,F., Cardozo,J.,  
Chambergo,F., Ciapina,L.P., Cicarelli,R.M.B., Coutinho,L.L.,  
Cursino-Santos,J.R., El-Dorriy,H., Faria,J.B., Ferreira,A.J.S.,  
Ferreira,R.C.C., Ferro,M.I.T., Formighieri,E.F., Franco,M.C.,  
Greggio,C.C., Gruber,A., Katsuyama,A.M., Kishi,L.T., Leite  
Jr.,R.P., Lemos,E.G.M., Lemos,M.V.F., Locali,E.C., Machado,M.A.,  
Madeira,A.M.B.N., Martinez-Rossi,N.M., Martins,E.C., Meidanis,J.,  
Mack,C.F.M., Miyaki,C.Y., Moon,D.H., Moreira,L.M., Novo,M.T.M.,  
Okura,V.K., Oliveira,M.C., Oliveira,V.R., Pereira Jr.,H.A., F.,  
Rossi,A., Sena,J.A.D., Silva,C., de Souza,R.F., Spinola,L.A.F.,  
Takita,M.A., Tamura,R.E., Teixeira,E.C., Tezza,R.I.D., Trindade dos  
Santos,M., Truffi,D., Tsai,S.M., White,F.F., Setubal,J.C. and  
Kitajima,J.P.

## TITLE

Comparison of the genomes of two Xanthomonas pathogens with  
differing host specificities

JOURNAL  
MEDLINE  
PUBMED

Nature 417 (6887), 459-463 (2002)  
22022145  
12024217

REFERENCE  
AUTHORS

2 (bases 1 to 10807)  
da Silva,A.C.R., Ferro,J.A., Reinach,F.C., Farah,C.S., Furlan,L.R.,  
Quaggio,R.B., Monteiro-Vitorello,C.B., Van Sluys,M.A., Almeida  
Jr.,N.F., Alves,L.M.C., do Amaral,A.M., Bertolini,M.C.,  
Camargo,L.E.A., Camarotte,G., Cannavan,F., Cardozo,J.,  
Chambergo,F., Ciapina,L.P., Cicarelli,R.M.B., Coutinho,L.L.,  
Cursino-Santos,J.R., El-Dorriy,H., Faria,J.B., Ferreira,A.J.S.,  
Ferreira,R.C.C., Ferro,M.I.T., Formighieri,E.F., Franco,M.C.,  
Greggio,C.C., Gruber,A., Katsuyama,A.M., Kishi,L.T., Leite  
Jr.,R.P., Lemos,E.G.M., Lemos,M.V.F., Locali,E.C., Machado,M.A.,  
Madeira,A.M.B.N., Martinez-Rossi,N.M., Martins,E.C., Meidanis,J.,  
Mack,C.F.M., Miyaki,C.Y., Moon,D.H., Moreira,L.M., Novo,M.T.M.,  
Okura,V.K., Oliveira,M.C., Oliveira,V.R., Pereira Jr.,H.A.,  
Rossi,A., Sena,J.A.D., Silva,C., de Souza,R.F., Spinola,L.A.F.,  
Takita,M.A., Tamura,R.E., Teixeira,E.C., Tezza,R.I.D., Trindade dos  
Santos,M., Truffi,D., Tsai,S.M., White,F.F., Setubal,J.C. and  
Kitajima,J.P.

TITLE  
JOURNAL

Direct Submission  
Submitted (28-NOV-2001) Departamento de Bioquimica, Universidade de  
Sao Paulo, Av. Prof. Lineu Prestes 748, Sao Paulo, SP 05508-900,  
Brazil

FEATURES  
source

Location/Qualifiers  
1. .10807  
/organism="Xanthomonas axonopodis pv. citri str. 306"  
/strain="306"  
/db\_xref="taxon:190486"  
/note="pathovar: citri"  
complement(149. .361)  
/gene="XAC1343"  
complement(149. .361)  
/gene="XAC1343"  
/note="putative; ORF located using Glimmer/Genemark"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="AAM36214.1"  
/db\_xref="GI:21107505"  
/translation="MPLKVDGTRIRPADGDPWEYVVVVVRNGKEIRRHVVDCAM  
HGGERGTLAVEASAVRSPGRTRH"  
complement(511. .1065)  
/gene="XAC1344"  
complement(511. .1065)  
/gene="XAC1344"  
/note="identified by sequence similarity; putative; ORF  
located using Blastx/Glimmer/Genemark"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"

## gene

complement(149. .361)  
/gene="XAC1343"  
complement(149. .361)  
/gene="XAC1343"  
/note="putative; ORF located using Glimmer/Genemark"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"

## CDS

complement(149. .361)  
/gene="XAC1343"  
complement(149. .361)  
/gene="XAC1343"  
/note="putative; ORF located using Glimmer/Genemark"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"

## gene

complement(511. .1065)  
/gene="XAC1344"  
complement(511. .1065)  
/gene="XAC1344"  
/note="identified by sequence similarity; putative; ORF  
located using Blastx/Glimmer/Genemark"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"

## CDS

complement(511. .1065)  
/gene="XAC1344"  
complement(511. .1065)  
/gene="XAC1344"  
/note="identified by sequence similarity; putative; ORF  
located using Blastx/Glimmer/Genemark"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"

/protein\_id="AAM36215.1"  
/db\_xref="GI:21107506"  
/translation="MSMRDQGNKKDGVPTPEVAADGRLFTAEASPAPVPAATP  
SVVPAAPTORHSEAKESLIAADISIEKGIEGAGHVRLAGRFKGDVNVKGLDITIERGA  
KMGDDGGKPVTTSKPWSGSDAT"  
complement(1315. .3024)  
/gene="XAC1345"  
complement(1315. .3024)  
/gene="XAC1345"  
/note="identified by sequence similarity; putative; ORF  
located using Blastx/Glimmer/Genemark"  
/codon\_start=1  
/transl\_table=11  
/product="sensor histidine kinase"  
/protein\_id="AAM36216.1"  
/db\_xref="GI:21107507"  
/translation="MSSSTFSRRRNQALIFSAIFVVGIGIFVGARRSLADAALVS  
HTHEVIGRVDEIQARLLDAESAQGFLLTNDAYLIGYQTSVERLPIILDNRRLIVD  
NPAQOHLAQLOLRVETLRQIQHVDIVTQSGLEPARADIQSAFRTTSAIRRLALA  
MYORELLAORAQSSQSANLLALALAGIPGLVVGMYALLMRLRHRAQAEOL  
AAQANRELGOSIDALQRTADNLNLSRYTGLLOSCISAEELMVSSTLASLLPGVAG  
SVLLRASODRAEAIHMGEPILVRSPHLSPEQWALRRGQPHIVDDUGDALCNSLR  
EPAHDASVTTACLPNSAQGTQGLFSLSAAGPGMPRLIEIAEAAAEQDLALSRLR  
ESLRQSRDALTGLYNNRYLEESLSHELARCARRGLPLVLMILDVDFHKQFNDSQGH  
AGDDLAAVGGELLRLRAEDVACRYGGEFTVLLPEADGGEAMRVAEQIRGYTAAL  
AVSDGARALPKVTASIGVASFPVDGEGASLIQKADAALYRAKHAGNRNVERHGAATD  
NAD"  
complement(3135. .3461)  
/gene="XAC1346"  
complement(3135. .3461)  
/gene="XAC1346"  
/note="identified by sequence similarity; putative; ORF  
located using Blastx/Glimmer/Genemark"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="AAM36217.1"  
/db\_xref="GI:21107508"  
/translation="WKTHMLSLSIAIALTLAACOGPDPAEQARRDAASAKVAAAATE  
QAADRAAEVKAATAADASVAARELAADASRATANAADATADTKREMANKAQNAKDHD  
ADHAQH"  
complement(3691. .4032)  
/gene="XAC1347"  
complement(3691. .4032)  
/gene="XAC1347"  
/note="identified by sequence similarity; putative; ORF  
located using Blastx/Glimmer/Genemark"  
/codon\_start=1  
/transl\_table=11  
/product="outer membrane protein"  
/protein\_id="AAM36218.1"  
/db\_xref="GI:21107509"  
/translation="MTINKLLIAMALGLALTACSKPEQAQDAASANEAAADAQTAAD  
QAASGSQSDAADAQAATAAASADTAADAQAASGAATDAADAQAADAQTAEDTAK  
AAKDTAEPAKK"  
complement(4155. .5330)  
/gene="atoB"  
/note="XAC1348"  
complement(4155. .5330)  
/gene="atoB"  
/note="identified by sequence similarity; putative; ORF  
located using Blastx/Glimmer/Genemark"  
/codon\_start=1  
/transl\_table=11  
/product="acetoacetyl-CoA thiolase"  
/protein\_id="AAM36219.1"  
/db\_xref="GI:21107510"  
/translation="MSDIVIVAARKTAIGSFQGNVAVPAPTLAAATQALAAQSGIA  
PADISEVIVGCVLPANLGPQAPRAAIAAGIPTGTINKVCSGKMALMGLHDLI  
KAGSASIVVAGMESNAPHLLPNRSRTGNFQAVDHMAWDGLTNPYDQAGMVF  
GEATAEKFGSRADODAFATASVRAHAORSGAFDEIVPVTVTRKGVVVDSEQ  
PKGSDVAKIPTLKPAPKDGSVTAASSSSISDGAATVVLMSADDAQRRGVTPLARIVG

gene	HVTHAQPEWFTTAPVAAIQSLVSKLWRLDDVDLFEINEAFVAVMAPIRQLGIAHD KVNVHGACALGHPIGASGARLVVTLVNALRSRGGKRGATLCTIGCGGEATAIALELI" 5447. .8326 /gene="XAC1349" CDS 5447. .8326 /gene="XAC1349" /note="Identified by sequence similarity: putative; ORF located using Blastx/Glimmer/Genemark" /codon_start=1 /transl_table=11 /product="serine protease" /protein_id="AAM36220.1" /db_xref="GI:21107511" /translation="MGPIHMGSGRTPMKKODVARTVLSALAVALSACGGGGGGG RSPTPAAPVTPSPPTSPPTTPPTPEPAIDAHALNTNRAAALGFTGAG YRIGVDTGINANHPALGRVSDFIYDPTNTNTAVGDVVGHTVVAELANGRAVQ WPGIAPGAGLVSAIISDRAPEDGGDNEVDPLGLGPVHADLLISAGVIMNNWS GGLFWDPTVNIQAIRYRPFILSHDGLVFAFSGNESRQSPDTPALPSQPPGNTLP AADLERGLVYVGDALDTPNTQLASYNACGVAMRYCLVAPGTSILFIDPDATAGNIRF YGSCTSTAPLVSCAALVQAPFYFNNDLVROTLLGTATDLAGAAGVDPVFGYGLLDV GKAVLGPARDMGTVVEDVTTLRSTWANDISGTGLTKRGTGALVLSSTANTFTGDQ VLGTQFLOSLOSRAVSIANGARLVGAGSGGRVDNAGTLQVNGATASAGNTQASN GLRANVYDRUNVAGTATAGDQLLQRRDYYVNNITYAVLQATNGLOGFTDLUSGP AVFLSATRSDYDANTAYVLSQMDVTAVALSLSGAIGTASMDAAIRVQOATQVDDAOF QNGALSGAETRAAAALQOSPQAKAADSLSLSGRAHDSAAFTFTIDLRRALAA HFDGLSQPRLGNMORALGPGEGGATTSGFATSGWMGNDLRLASGATVGFAGFT RSSLSGLDGLARGDROVQALYWGTTLGAATLGGONGFNVNRQIERTLQLGDRSS AFSDYSYSSLSNLETGYRLGGAHSLTPYMGLDYVRLSRDGFRESGGDFGLRANAS TSSQALGAVRSYAHWRGQIOLGGYAEWQALSSDGLTLDAFVGVDAWAPLRFQWQA RSGGLGFTAAPIGQSQSRFQYDQVRGGRGDDRLSLRSADF" complement(8388. .9488) /gene="leu" /note="XAC1350" complement(8388. .9488) /gene="leu" /note="Identified by sequence similarity: putative; ORF located using Blastx/Glimmer/Genemark" /codon_start=1 /transl_table=11 /product="leucine dehydrogenase" /protein_id="AAM36221.1" /db_xref="GI:21107512" /translation="MFLFETLDTTGHQVIFCHNRDAGLKAIIALHSTRLGPALGGVRM RPYASSFAALNDALRLTMTYTNKALAGNVGGKAVIIGDPRDKTEALFRAGREV DTLGRITSDVDVDMEOIYLESEYVTGVHVRHSGDPAPPTAYGALQALMAS LNRKLHGEVKGASIAIOLGHMELVKLLKRGAKLYVADLNQALVDRAVAYGAE AVRDEIHQVSADVFAPCALEAGNEDTLPLQKAKIICGTANNQLASAAAGEELHRG
gene	
CDS	
Query Match	2.4%; Score 77.4; DB 1; Length 10807;
Best Local Similarity	52.7%; Pred. No. 7.4e-07;
Matches	246; Conservative 0; Mismatches 206; Indels 15; Gaps 3;
QY 1224	CGGAACATCCITTTCCGCACCCATCGTACCGGCACGGCGCTCTGCTGCGCAGAAATA 1283
DB 6457	CGGCACCTCGTTCGCAGCGCGCTGCTCCGGTCCGCGCGCTGGTGTGGCAGGCATT 6516
QY 1284	CCCGTGGATGAGCAACAACTCCGTACCAACGCTGCTGCAACGGCTCAGGACATCGG 1343
DB 6517	TCCGTATTTCAACATGATCTGGTGGCGCCAGCACCTTGTGGGAATGCCACGACTTGG 6576
QY 1344	TGCACTCGCGCTGGACAGCAAGTTCCGCTGGGACTGCTGATCGCGGTAAAGCCATGAA 1403
DB 6577	GSCCGCGGTGTCGACCGGTTATTCGGCTACGGCTGCTGACGCTCGCAAGCGCGTGTCT 6636
QY 1404	CGGACCCCGCTCTTCCTTCGCTGGGACTTTTACCCCGCATAGGAAGGTACATCCGATAT 1463
DB 6637	GGGCGCGCAGCTTTCGATTTGGGCGCAGCTGGAGTCTCATGTCA-----CCACCTT 6687
QY 1464	TGCTTACTCTCCCTGAACGACATTTTCAGGACCGCGCGCTGATCAAAAAGGCGGAC 1523
DB 6688	GGTTTCGACCTGGGCGCACGACATCAGTGGCACCGCGGCTGACCAACGCGGACCGG 6747
QY 1524	CCAACTGCAACTGCACGGCA---ACAACACCTTATACGGGCAAAACCATTTATCGAAGCGG 1580

Db 6748	CGCGTGTGTTATGAGCAGCAGCGGCAATACCTTACCAGCGCATACCCAGTACTCGCGG 6807
QY 1581	TTCGCTGCTGTTGTACGCAACAACAATCGGATATGCGCTCGAAACCAAGGTGCGCT 1640
Db 6808	CACCTTGCAGACCGCGCATGTCGAGAGTGCAGCGGTGAGCATCG---CCAACGCGCGCG 6864
QY 1641	CATTATATAACCGCGGCATCCGCGGTAGCTGAACAGCAGCAGCGCA 1687
Db 6865	CCTGTCGCGCGCGGCACATCGGTGCTCGGTCGCAACACGCGGGA 6911
RESULT 31	
AE012229	
LOCUS	AE012229 10029 bp DNA linear BCT 23-MAY-2002
DEFINITION	Xanthomonas campestris pv. campestris str. ATCC 33913, section 137 of 460 of the complete genome.
ACCESSION	AE012229 AE008922
VERSION	AE012229.1 GI:21112347
KEYWORDS	Xanthomonas campestris pv. campestris str. ATCC 33913.
SOURCE	Xanthomonas campestris pv. campestris str. ATCC 33913.
ORGANISM	Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group; Xanthomonas.
REFERENCE	1 (bases 1 to 10029)
AUTHORS	da Silva,A.C.R., Ferro,J.A., Reinach,F.C., Farah,C.S., Furlan,L.R., Quaggio,R.B., Monteiro-Vitorello,C.B., Van Sluys,M.A., Almeida Jr.,N.F., Alves,L.M.C., do Amaral,A.M., Bertolini,M.C., Camargo,F., Ciapina,L.P., Ciccarelli,R.M.B., Coutinho,L.L., Churino-Santos,J.R., El-Dorri,H., Faria,J.B., Ferreira,A.J.S., Cursino-Santos,J.R., El-Dorri,H., Faria,J.B., Ferreira,A.J.S., Ferreira,R.C.C., Ferro,M.I.T., Formighieri,E.F., Franco,M.C., Greggio,C.C., Gruber,A., Katsuyama,A.M., Kishi,L.T., Leite Jr.,R.P., Lemos,E.G.M., Lemos,M.V.F., Locali,E.C., Machado,M.A., Madeira,A.M.B.N., Martinez-Rossi,N.M., Martins,E.C., Meidanis,J., Menck,C.F.M., Miyaki,C.Y., Moon,D.H., Moreira,L.M., Novo,M.T.M., Okura,V.K., Oliveira,M.C., Oliveira,V.R., Pereira Jr.,H.A., Rossi,A., Sena,J.A.D., Silva,C., de Souza,R.F., Spinola,L.A.F., Takita,M.A., Tamura,R.E., Teixeira,E.C., Tezza,R.I.D., Trindade dos Santos,M., Truffi,D., Tsai,S.M., White,F.F., Setubal,J.C. and Kitajima,J.P.
TITLE	Comparison of the genomes of two Xanthomonas pathogens with differing host specificities
JOURNAL	Nature 417 (6887), 459-463 (2002)
MEDLINE	22022145
PUBMED	12024217
REFERENCE	2 (bases 1 to 10029)
AUTHORS	da Silva,A.C.R., Ferro,J.A., Reinach,F.C., Farah,C.S., Furlan,L.R., Quaggio,R.B., Monteiro-Vitorello,C.B., Van Sluys,M.A., Almeida Jr.,N.F., Alves,L.M.C., do Amaral,A.M., Bertolini,M.C., Camargo,L.E.A., Camarotte,G., Cannavan,F., Cardoso,J., Churino-Santos,J.R., El-Dorri,H., Faria,J.B., Ferreira,A.J.S., Ferreira,R.C.C., Ferro,M.I.T., Formighieri,E.F., Franco,M.C., Greggio,C.C., Gruber,A., Katsuyama,A.M., Kishi,L.T., Leite Jr.,R.P., Lemos,E.G.M., Lemos,M.V.F., Locali,E.C., Machado,M.A., Madeira,A.M.B.N., Martinez-Rossi,N.M., Martins,E.C., Meidanis,J., Menck,C.F.M., Miyaki,C.Y., Moon,D.H., Moreira,L.M., Novo,M.T.M., Okura,V.K., Oliveira,M.C., Oliveira,V.R., Pereira Jr.,H.A., Rossi,A., Sena,J.A.D., Silva,C., de Souza,R.F., Spinola,L.A.F., Takita,M.A., Tamura,R.E., Teixeira,E.C., Tezza,R.I.D., Trindade dos Santos,M., Truffi,D., Tsai,S.M., White,F.F., Setubal,J.C. and Kitajima,J.P.
TITLE	Direct Submission
JOURNAL	Submitted (28-NOV-2001) Departamento de Bioquimica, Universidade de Sao Paulo, Av. Prof. Lineu Prestes 748, Sao Paulo, SP 05508-900, Brazil
FEATURES	Location/Qualifiers 1. .10029 /organism="Xanthomonas campestris pv. campestris str. ATCC 33913" /strain="ATCC 33913" /db_xref="ATCC:33913"









```
-10_signal 59. .64
-35_signal 164. .169
-10_signal 187. .192
RBS 243. .249
gene 257. .1705
CDS /gene="aprPF33"
257. .1705
/gene="aprPF33"
/codon_start=1
/transl_table=11
/product="alkaline protease"
/protein_id="BAA36461.1"
/db_xref="GI:4115623"
/translation="MSKVKDKAIYSAQAQSTAYTQIDSFSHQYDRGNLTVNGKPSFT
VDAQADHLRENAAYRDVGNKIDLTFTLTSTASNATMKNKHGITGFQSFNTQOKAQA
VLAQSNADVANVTFEKASGDFHMTFGNYSGGQDGAAPAYLPGTNAKYNESGLDG
TSWLTNSYTPNKTPLDNNYGRQTLTHEIGHTLGLDHPGDYNAGTGNPSYKDDYGO
DTRGYSNYSSESTNWFSGKVEAYSSGLDIDDAIOKLYGANYNTRAGDTYTG
FNSNTGRDLSATSSADKLVSFVWDGGGNDTLDPSGTQOKINLNETSFSDVGGSLYG
NVSTAKGVTYVNAFGSGNDLLIGNAANILKGGAGNDIYGGGGADQLMGAGSDTF
VFGSSDSKPCAADKIFDFTSGSDKIDLTTGTTKGAGLTFVNAFTGHAGDAVLSYASGT
NLCTLAVDFSGHGVADFLVTVGQAAYSDIVA"
1765. .1767
RBS 1774. .2148
gene /gene="inhPF33"
1774. .2148
CDS /gene="inhPF33"
/feature="note" note="putative"
/codon_start=1
/transl_table=11
/product="protease inhibitor"
/protein_id="BAA36462.1"
/db_xref="GI:4115624"
/translation="MVMTLFTTYVWCFOAMLMSAGASAMASSLVLPISAGLAGHWAL
HQEQVCALDIEQANLGGDVACAGQWLGEAPVSWPTPDGILWLNAGTGGTITHLNR
QKEGEYGRSGSGLVLLQRAP"
2236. .2243
RBS 2253. .4004
gene /gene="aprDPF33"
2253. .4004
CDS /gene="aprDPF33"
/function="ABC exporter"
/codon_start=1
/transl_table=11
/product="ABC protein"
/protein_id="BAA36463.1"
/db_xref="GI:4115625"
/translation="MAKTPVPAPLFKALGDYKSIILISIGCTFALINVLMLVPSIYMLQ
VYDRLVSSQNETTLVMTLTMVYGFPAFVIGVLEIRSFIVIRIGSOLERRERNLRYKKA
FERNLQREGHGSGSLGDLTHIQFITGPALFADFAPWFTIYLFVLFVWLGVLGA
SAGAVLLTGLACLNXYLTKKPLGEASGYSOOSTOLATSHLNAETIOAMGLALKRK
WFKVHSOPLGLQNKASDTGSIISLSKSLRLCLOSLVGLGALLVIKGDWTAGMWIAG
SILMGVSLPDLQILAVYWNQWSAKLAYORLDNLMRPPPOGEQOMALPAPKGNVSFEQ
VSAGPPGRVPTLHVQSNLAAGEVLVGLGASGSKSTLARVLGVVWPTLAGTVRLDG
ADIHRWDRDLGLGHIGLPQDIELFSGSIADNIARFRDADPOLVYQAAQAGVHLIL
RLPHGYDTVLDEGSLSGQKQKRALARALYGGPRLIVLDPEPNSLDTVGEAALASA
IMQMAQGSTVVLVYTHRSSALAQDKLLVLNNEGRLOAFGPPSQDVLRLALSGQOEAPEK
AGVLSLSQYQAGRPGA"
3986. .3991
RBS /gene="aprDPF33"
4001. .5335
gene /gene="aprPF33"
4001. .5335
CDS /gene="aprPF33"
/function="ABC exporter (MFP)"
/codon_start=1
/transl_table=11
/product="Membrane fusion protein (MFP)"
/protein_id="BAA36464.1"
/db_xref="GI:4115626"
/translation="MSDSSSIQINRDYEQTHPEHGAREFFARMGNLLTVVAGGFFLWA
SLAPUDQIPYQGVTVVYSGKKRAQNTLSPGVVSRLVREGEAVKQGGPFLRDTQNG
ADVHSLQAOYRMAVASVARMQSERDNQSTITTFPAELSGNPQDALALVLEGORQLFSR
```

```
REAPAREQAQIRANIEGATAQLNGMRRRSDLTAAQASLRDQLNNLQPLADNGYIPRN
RLMEYQRLSQVQODLAQNTGESGRVQGLESLRLKLOHQHSEYEQKVRSLQADQLR
SLTLEQOLTSAQFDLOHSEINAPADGIAVNLVSVHTEGAVRAGETLLEIVPODTRLEV
EGRLPVHLVDKVGTHLPVDILFTAFNOSRTPRVPCEVSLISADQMLDEKTMPIYVLR
TTVSSSALXKLHGLVIRPGMAEMFIRTGERSLLNLYLFKPLLDAGSALTEE"
5327. .5332
RBS /gene="aprEPF33"
5338. .6672
gene /gene="aprPF33"
5338. .6672
CDS /gene="aprPF33"
/function="ABC exporter (OMP)"
/codon_start=1
/transl_table=11
/product="outer membrane protein (OMP)"
/protein_id="BAA36465.1"
/db_xref="GI:4115627"
/translation="MKPMFIALIACSSAAQAMGPFVYEOALRNDPVFLGAIKERDA
GLENRTIGRAGLLPKLSYNYNKGNNNSQATLPDGRGNYHDDRNYNSGYSTFTLQQL
FDYAYANYRKGVAQALFADSFDRDKSOALLVRLVITYTOALFAQDQIDITARAKKAY
EOQFOQNHLEFOQEGTTRTDILEAESVELATAEIOALDEQDASLRELGAJGVQSV
NINDLALNPGCFAPATLTPANYDTHWELALTNNPTLASORQSLEVAEYVERNEAGHM
PRVTAYASSRQSSDSGNTNQRDNTNGVEISIMPLYAGGVSASTRQASRAMQEAE
YELEGKTRETLIELRRQFSACLGVSKLRYQKALTSAEALVSTKQILGGERKNLD
ALNAEQOYLSTRDLAQARYDYLANMTKLHYAGNLRDRTLAKYDEAFGPEH"
6692. .6697
RBS 6706. .9663
gene /gene="pspA"
6706. .9663
CDS /gene="pspA"
/codon_start=1
/transl_table=11
/product="serine protease homologue"
/protein_id="BAA36466.1"
/db_xref="GI:4115628"
/translation="MITDSPRFKPFTAGSLLLLLSVAAHGOYLEAGKPGDPASMRSTEF
QSDWGLDRMKANEAYAGITGSGYKIGALDSGDFDPNHPKASKDRHYHATVAGTYVDS
PFNTGALNPNDNSHGTHVTGTMGAARDGVMHGVAYNAQIYVGNTNANDSFLPGPTP
DPQYFKAVYSALVDSGVRAINNSWGSDQKDVSYOTLDGLHAAAYAOHYNONTWLDAAAD
VAKAGVINVSAGNSGYANASVRSALPYFQPLEGHVLAVSGLDKANNOKYNOCGIK
WYCLSTPGALINSTIPEGGYGVKSGTSMSPHATGALALMYERYPYLLNNEQALQVLIT
TATQNLGAVTADPTQVGVGPDILGRAMHGPQLLGPMTVTLAAGQDQVSNALSDKA
LIRQEAEDRAETLAWQOTLRDKWENGWVGAGSQDQDQDYAVGTARAAAQAQRIYEGS
LTKSGAGSLILTGDNTRYGPTVNGSLVNGSLTSATVYDNGSLTGGSGRTAALQVN
NGGTVAPGNSVGTQVAGDVNLAAGSYAVELTPRSQIVATGKATLGGTAVTLALE
NSPTLLSQOQAESLVRQYTTIOAAGGIIQOGFAGVLPNLYFLGCTLDYAAAGVOLATA
RSDASFASVGNTRNORSVAAAEOQLAGNPVYESILRSDSLATAQOQFOOLSGEYPA
IGTMLINDSROVDAGVERLHRTVPAGESNVVWKAALSGWGTDSRSETAGTSSILGM
LAGVDGALDDQOTRVGLVAGYSDSLNMGSGTHSSASIDSYHLGAYAGHELGNWRLSVG
GAYSWHRGDYKRDLOQYGEVSGKQKARLDAKSAQVTEAAAYRLNYPVALEFPANLAYV
HLDNERPHEKGDAAAALARGSDSDAVLSTLGLRASKTALANDROQLDVSIGSLGWOHSL
SAVSEBHLAFVAGGSPFAVQSSPLMRDAALVGVQASLALSLSKSTRVNLDTYTGQLAGRE
KIQGVGLSLNQF"
9835. .9842
RBS /gene="pspB"
9848. .12958
gene /gene="pspB"
9848. .12958
CDS /gene="pspB"
/codon_start=1
/transl_table=11
/product="serine protease homologue"
/protein_id="BAA36467.1"
/db_xref="GI:4115629"
/translation="MKTTRKGFOSAIPGCGYPLKALSGVPYGAALCCLASLGTAAQA
PYVETGKLGDAAASWRSNEFKADWGLGAVHADAAAYAGYTGKGVKLGIFDQVYQAQHE
FASPQKVVTIVTEIGRIQTDPIPVKAGDAPYDGTFTSLGSLGNHNGHVVGGIAG
NRDGGPMHGVAFDAQIIISAENDPGPEDGIILGNDGAVYKAGWDALVAGARIINNSW
GIGIGDQYAKGGRDPAPFNFTVNEAQOFNNIRPIILGTIAGGATOGAIDAAASGVLLTI
FAACNDYNRRNNPDAISGLAYFVPIAPNPLVSAALOQNPDFTASADPVYISTFSRCGY
AASFCAPACTKIIYSSVYNTGTLLENLTDDYANFNCTMAAPHPVAGSAALVMEFPYMS
CDQJSTLLKTTATDLGAPGIDSLYGMWNLGKALINGPFGMEVTAEDIPAERIDGAYG
SGQVADLPGMGAVVDAGKTPQRCTDTHCCGLDWSNDIAGHGLGTLKQIGLVALITGT
NTYSGPTLVNQGLLAINGSTSDVTYVGSSGVVGGSGHIGTSLYANSNGGTVPAGNSIGTL
```





FEATURES	13083-970, Brazil	Location/Qualifiers	AATLFATALLASGINSVTATLAGQIVMEGFLRLRLPWLRRVLTGRGLAIVPVVVVA LYXEQGTGRLLLSQVILSQMLPFAVILPLRLCVRADKVMGALVAPRWLWVWMLIAGV IVLVNKLGLDVAVHLMWGVSD"	
	source			
gene	1. .11186	/organism="Xylella fastidiosa 9a5c" /db_xref="taxon:160492" /clone="9a5c" complement(128. .631) /gene="XF1012" complement(128. .631) /gene="XF1012" /note="similar to SP P30134 (percent identity: 54 %/query alignment coverage: 88.6 %/subject alignment coverage: 83.1 %); identified by sequence similarity; putative; ORF located using Glimmer/RBSfinder"	2720. .2884 /gene="XF1016" 2720. .2884 /gene="XF1016" /note="hypothetical protein; identified by sequence similarity; putative; ORF located using Glimmer/RBSfinder" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="AAF83826.1" /db_xref="GI:9105954" /translation="MIRQRCQSQWSMCFVLMHRFAVFOGAVYMPVCMVVDGSSRRHN DVLAICWVSF" complement(2862. .3035) /gene="XF1017" complement(2862. .3035) /gene="XF1017" /note="hypothetical protein; identified by sequence similarity; putative; ORF located using Glimmer/RBSfinder" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="AAF83827.1" /db_xref="GI:9105955" /translation="MNRTRKDSSTTNELFQPPQPNLTTSQNTIELSAPAPQPKQTAS WHDASHIKTIPNK" complement(3202. .3966) /gene="XF1018" complement(3202. .3966) /gene="XF1018" /note="similar to GI 6321421 (percent identity: 25 %/query alignment coverage: 95.7 %/subject alignment coverage: 48.3 %); identified by sequence similarity; putative; ORF located using Glimmer/RBSfinder/Start codon shift: -99" /codon_start=1 /transl_table=11 /product="arginine-tRNA-protein transferase" /protein_id="AAF83828.1" /db_xref="GI:9105956" /translation="MAIDSKPHDDQLQFLTNLHPCGYWPDWASDLVMDPNDRPLGAI YQALAWGFRSRGNLLYRPHCEHCNACVPVRVNVAFVFNRRSQRCLARNATLVTRIV PARNAEQLSLYRRYLHQRPDGMGHAIEFDQFLIGPWGYGRFMEIREPATNCTP GOLLAVATDLTHOALSAYVTFEVPNAAAGLGTALILQIHWAOERQRPYLVLGYWI KDFKMDYKRRFKLEIYDGYRWRPFSTTYPPTHLL" 4148. .4267 /gene="XF1019" 4148. .4267 /gene="XF1019" /note="hypothetical protein; identified by sequence similarity; putative; ORF located using Glimmer/RBSfinder" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="AAF83829.1" /db_xref="GI:9105957" /translation="MLRAITKGLFALLGVNAGFVVVLGQVCGEGVLDVWVWMLL" complement(4388. .4879) /gene="XF1020" complement(4388. .4879) /gene="XF1020" /note="similar to GI 155377 (percent identity: 48 %/query alignment coverage: 101.8 %/subject alignment coverage: 98.8 %); identified by sequence similarity; putative; ORF located using Glimmer/RBSfinder" /codon_start=1 /transl_table=11 /product="pathogenicity-related protein"	
	CDS			
gene	1102. .1284	/product="conserved hypothetical protein" /protein_id="AAF83822.1" /db_xref="GI:9105950" /translation="MTATPPPPALPEHIHMRHALTLAHRATEFDEIPVGAVLISPE GTLLEGWNYNITSDHPSAHEAIEIMALRAAGHOLNRHMPGCTLYVLEPCLMACMAII HARIAHLIYAAADPKTGACGSTFDIINDPRHHNHVHVYVGLLAEEASRLTNYFRTRK GQPLPQ" complement(639. .1121) /gene="XF1013" complement(639. .1121) /gene="XF1013" /note="similar to GI 1787038 (percent identity: 61 %/query alignment coverage: 90.0 %/subject alignment coverage: 92.9 %); identified by sequence similarity; putative; ORF located using Glimmer/RBSfinder/Start codon shift: 12" /codon_start=1 /transl_table=11 /product="conserved hypothetical protein" /protein_id="AAF83823.1" /db_xref="GI:9105951" /translation="MNRHTIPKHTQQTAPLIDAKIHVASFRQVRQARRDELKID YVELISDLSAQARQVDIAAHLGLVAOPTAKMLKRLARDGWIQIRPYRGTFLLPAG KALAAASRECHQTVRFLLALGIDAATARRDAEGIEHHVSETTLAAFRALHRRHSDTT " 1102. .1284 /gene="XF1014" 1102. .1284 /gene="XF1014" /note="hypothetical protein; identified by sequence similarity; putative; ORF located using Glimmer/RBSfinder/Start codon shift: -12" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="AAF83824.1" /db_xref="GI:9105952" /translation="MWCLVFICGICVGHRYFGVTFLCLRPSHGESALGVDPWDVLNGEM LVSCVWLCLWAAIFL" 1286. .2704 /gene="XF1015" 1286. .2704 /gene="XF1015" /note="similar to GI 5670182 (percent identity: 64 %/query alignment coverage: 93.6 %/subject alignment coverage: 100.9 %); identified by sequence similarity; putative; ORF located using Glimmer/RBSfinder" /codon_start=1 /transl_table=11 /product="manganese transport protein" /protein_id="AAF83825.1" /db_xref="GI:9105953" /translation="MSCGIANAIRWDVSLIHRSSSSMSVQPTSPVSDSPSLGEMH ASVAVSRRCWHGFRLLAFILPGYVMSVGMPCGNWATGLAGGSRFCYMLLSVLLSNV MAIVLQALRAGIASDMDLAQACRARSRGTTLAWVVCELAIACDLAEVIGTIAIA LMLLGVPIWGQVITAVDVLVLLMLHRRGFRALFAFVITALLVIFGCRVQVILAAP PIQEVLGVPQWQVADPOALYLAIGIVGATVMPHNLHSSIIVQTRAYPTPVGRR SALRWAVADSTLALMLF INASILIILAAVFAQHFDVEIEQAYOLLAPVLGVGV	2720. .2884 /gene="XF1016" 2720. .2884 /gene="XF1016" /note="hypothetical protein; identified by sequence similarity; putative; ORF located using Glimmer/RBSfinder" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="AAF83826.1" /db_xref="GI:9105954" /translation="MIRQRCQSQWSMCFVLMHRFAVFOGAVYMPVCMVVDGSSRRHN DVLAICWVSF" complement(2862. .3035) /gene="XF1017" complement(2862. .3035) /gene="XF1017" /note="hypothetical protein; identified by sequence similarity; putative; ORF located using Glimmer/RBSfinder" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="AAF83827.1" /db_xref="GI:9105955" /translation="MNRTRKDSSTTNELFQPPQPNLTTSQNTIELSAPAPQPKQTAS WHDASHIKTIPNK" complement(3202. .3966) /gene="XF1018" complement(3202. .3966) /gene="XF1018" /note="similar to GI 6321421 (percent identity: 25 %/query alignment coverage: 95.7 %/subject alignment coverage: 48.3 %); identified by sequence similarity; putative; ORF located using Glimmer/RBSfinder/Start codon shift: -99" /codon_start=1 /transl_table=11 /product="arginine-tRNA-protein transferase" /protein_id="AAF83828.1" /db_xref="GI:9105956" /translation="MAIDSKPHDDQLQFLTNLHPCGYWPDWASDLVMDPNDRPLGAI YQALAWGFRSRGNLLYRPHCEHCNACVPVRVNVAFVFNRRSQRCLARNATLVTRIV PARNAEQLSLYRRYLHQRPDGMGHAIEFDQFLIGPWGYGRFMEIREPATNCTP GOLLAVATDLTHOALSAYVTFEVPNAAAGLGTALILQIHWAOERQRPYLVLGYWI KDFKMDYKRRFKLEIYDGYRWRPFSTTYPPTHLL" 4148. .4267 /gene="XF1019" 4148. .4267 /gene="XF1019" /note="hypothetical protein; identified by sequence similarity; putative; ORF located using Glimmer/RBSfinder" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="AAF83829.1" /db_xref="GI:9105957" /translation="MLRAITKGLFALLGVNAGFVVVLGQVCGEGVLDVWVWMLL" complement(4388. .4879) /gene="XF1020" complement(4388. .4879) /gene="XF1020" /note="similar to GI 155377 (percent identity: 48 %/query alignment coverage: 101.8 %/subject alignment coverage: 98.8 %); identified by sequence similarity; putative; ORF located using Glimmer/RBSfinder" /codon_start=1 /transl_table=11 /product="pathogenicity-related protein"	
	CDS			
gene	1286. .2704	/product="manganese transport protein" /protein_id="AAF83825.1" /db_xref="GI:9105953" /translation="MSCGIANAIRWDVSLIHRSSSSMSVQPTSPVSDSPSLGEMH ASVAVSRRCWHGFRLLAFILPGYVMSVGMPCGNWATGLAGGSRFCYMLLSVLLSNV MAIVLQALRAGIASDMDLAQACRARSRGTTLAWVVCELAIACDLAEVIGTIAIA LMLLGVPIWGQVITAVDVLVLLMLHRRGFRALFAFVITALLVIFGCRVQVILAAP PIQEVLGVPQWQVADPOALYLAIGIVGATVMPHNLHSSIIVQTRAYPTPVGRR SALRWAVADSTLALMLF INASILIILAAVFAQHFDVEIEQAYOLLAPVLGVGV	2720. .2884 /gene="XF1016" 2720. .2884 /gene="XF1016" /note="hypothetical protein; identified by sequence similarity; putative; ORF located using Glimmer/RBSfinder" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="AAF83826.1" /db_xref="GI:9105954" /translation="MIRQRCQSQWSMCFVLMHRFAVFOGAVYMPVCMVVDGSSRRHN DVLAICWVSF" complement(2862. .3035) /gene="XF1017" complement(2862. .3035) /gene="XF1017" /note="hypothetical protein; identified by sequence similarity; putative; ORF located using Glimmer/RBSfinder" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="AAF83827.1" /db_xref="GI:9105955" /translation="MNRTRKDSSTTNELFQPPQPNLTTSQNTIELSAPAPQPKQTAS WHDASHIKTIPNK" complement(3202. .3966) /gene="XF1018" complement(3202. .3966) /gene="XF1018" /note="similar to GI 6321421 (percent identity: 25 %/query alignment coverage: 95.7 %/subject alignment coverage: 48.3 %); identified by sequence similarity; putative; ORF located using Glimmer/RBSfinder/Start codon shift: -99" /codon_start=1 /transl_table=11 /product="arginine-tRNA-protein transferase" /protein_id="AAF83828.1" /db_xref="GI:9105956" /translation="MAIDSKPHDDQLQFLTNLHPCGYWPDWASDLVMDPNDRPLGAI YQALAWGFRSRGNLLYRPHCEHCNACVPVRVNVAFVFNRRSQRCLARNATLVTRIV PARNAEQLSLYRRYLHQRPDGMGHAIEFDQFLIGPWGYGRFMEIREPATNCTP GOLLAVATDLTHOALSAYVTFEVPNAAAGLGTALILQIHWAOERQRPYLVLGYWI KDFKMDYKRRFKLEIYDGYRWRPFSTTYPPTHLL" 4148. .4267 /gene="XF1019" 4148. .4267 /gene="XF1019" /note="hypothetical protein; identified by sequence similarity; putative; ORF located using Glimmer/RBSfinder" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="AAF83829.1" /db_xref="GI:9105957" /translation="MLRAITKGLFALLGVNAGFVVVLGQVCGEGVLDVWVWMLL" complement(4388. .4879) /gene="XF1020" complement(4388. .4879) /gene="XF1020" /note="similar to GI 155377 (percent identity: 48 %/query alignment coverage: 101.8 %/subject alignment coverage: 98.8 %); identified by sequence similarity; putative; ORF located using Glimmer/RBSfinder" /codon_start=1 /transl_table=11 /product="pathogenicity-related protein"	2720. .2884 /gene="XF1016" 2720. .2884 /gene="XF1016" /note="hypothetical protein; identified by sequence similarity; putative; ORF located using Glimmer/RBSfinder" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="AAF83826.1" /db_xref="GI:9105954" /translation="MIRQRCQSQWSMCFVLMHRFAVFOGAVYMPVCMVVDGSSRRHN DVLAICWVSF" complement(2862. .3035) /gene="XF1017" complement(2862. .3035) /gene="XF1017" /note="hypothetical protein; identified by sequence similarity; putative; ORF located using Glimmer/RBSfinder" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="AAF83827.1" /db_xref="GI:9105955" /translation="MNRTRKDSSTTNELFQPPQPNLTTSQNTIELSAPAPQPKQTAS WHDASHIKTIPNK" complement(3202. .3966) /gene="XF1018" complement(3202. .3966) /gene="XF1018" /note="similar to GI 6321421 (percent identity: 25 %/query alignment coverage: 95.7 %/subject alignment coverage: 48.3 %); identified by sequence similarity; putative; ORF located using Glimmer/RBSfinder/Start codon shift: -99" /codon_start=1 /transl_table=11 /product="arginine-tRNA-protein transferase" /protein_id="AAF83828.1" /db_xref="GI:9105956" /translation="MAIDSKPHDDQLQFLTNLHPCGYWPDWASDLVMDPNDRPLGAI YQALAWGFRSRGNLLYRPHCEHCNACVPVRVNVAFVFNRRSQRCLARNATLVTRIV PARNAEQLSLYRRYLHQRPDGMGHAIEFDQFLIGPWGYGRFMEIREPATNCTP GOLLAVATDLTHOALSAYVTFEVPNAAAGLGTALILQIHWAOERQRPYLVLGYWI KDFKMDYKRRFKLEIYDGYRWRPFSTTYPPTHLL" 4148. .4267 /gene="XF1019" 4148. .4267 /gene="XF1019" /note="hypothetical protein; identified by sequence similarity; putative; ORF located using Glimmer/RBSfinder" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="AAF83829.1" /db_xref="GI:9105957" /translation="MLRAITKGLFALLGVNAGFVVVLGQVCGEGVLDVWVWMLL" complement(4388. .4879) /gene="XF1020" complement(4388. .4879) /gene="XF1020" /note="similar to GI 155377 (percent identity: 48 %/query alignment coverage: 101.8 %/subject alignment coverage: 98.8 %); identified by sequence similarity; putative; ORF located using Glimmer/RBSfinder" /codon_start=1 /transl_table=11 /product="pathogenicity-related protein"
	CDS			

Query Match

Best Local Similarity 2.2%; Score 71.2; DB 1; Length 11186;

51.2%; Pred. No. 2e-05;

Query Match 2.2%; Score 71.2; DB 1; Length 11186;  
Best Local Similarity 51.2%; Pred. No. 2e-05;



```
CDS
1431..3092
/genes="RSc0004"
/function="miscellaneous; hypothetical/global homology"
/note="Product confidence : probable
Gene name confidence : hypothetical
Predicted by Codon_usage
Predicted by Homology
Predicted by Framed"
/codon_start=1
/evidence-not_experimental
/transl_table=11
/product="PROBABLE TRANSMEMBRANE PROTEIN"
/protein_id="CAD13532.1"
/db_xref="GI:17427012"
/translation="MDIKRTILWVIFSLVLLFDNMORANGHOSFFPTQVTTTAA
AAGGCTPAGDVPKAAAPAAAGSOAPATGAVSQTPASEKLVTVTDVIRATVPAGAI
VKLELLTQDHGNGPNVLFDRSLERTYLARSLGIGDFPNHTTVFTASAGPRDLGTGG
EVSLLTQDHGNAKLAITYFKRGSYVIDTREFDNDGAAPINPTLYMELARDGAVE
QSRFSTFFGVDRIDTDFHYKITEFADIDKSAHVPAPTDSGVWAVQHYFASAWIPA
ASAKREFFYDRIIDTFRVGMQALGTVPAGSATSATRLFAGPQERMLEGITPGLE
LVKDYGLWTIIAKPLFWLLEKTHKLGNGWSIVALTVLKLVFFPLSATSVRSMAKM
KLDQPMATIRHKGDPQKMQNMWTLYTEKVNPLGCLPIVIOIPFIALLYVLL
SSVMRGAPWLGWHDLASDPFIILPILMAVSMFVQTRLNPTPPDPVQAKMMFMPI
AFVSMFFFPGLVLYVWVYVNNCLSIQAQWSINRMGLTNKKAATAA"
3222..4667
/genes="thdF"
/note="RSc0005; RS01827"
3222..4667
/genes="thdF"
/function="cell processes; protection responses;
detoxification"
/note="Product confidence : probable
Gene name confidence : probable
Predicted by Codon_usage
Predicted by Homology
Predicted by Framed"
/codon_start=1
/evidence-not_experimental
/transl_table=11
/product="PROBABLE THIOPHENE AND FURAN OXIDATION PROTEIN"
/protein_id="CAD13533.1"
/db_xref="GI:17427013"
/translation="MTASHAMTSPTVSDAASAPIRTVPAAIATAPRGGGIGVVRVS
GPDVAVMAVGCRLLPPQATYLPFLDAGAIIDRGAIALFWFPAPHSYGEDVLELQG
HGGPVYMLQSLRAGHIGLURVAEPGETFRAPLNDKLDLAQAEAVADILEASTE
AAARSAVSLSDGVSFTVHALVERVHLRMLVEATLDFPEEIDFLEAADARQLADI
RARLDVLAQRGALLREGLHVVLAQPNVGSLLNALAGAEALVTPITAGTTRDK
VOOTIOIEGIPUNIVDTAGTDEVERIGERTWAAIARADVVLHLILDADYRAGH
LSAEDAAIDARTAEHVPVPTLRVINKIDLAGAAYPGRVDAQPPPEWLSARDGIE
LIRAALETAGVGGEGLYLARERHLSALRSAREHLTTAAADHADQRAQSLDLFAEEL
RLAQEALNSITGAFSSDDLGLGVFSRFCIGK"
4757..5080
/genes="RSc0006"
/note="RS01828"
4757..5080
/genes="RSc0006"
/function="miscellaneous; hypothetical/global homology"
/note="Product confidence : hypothetical
Gene name confidence : hypothetical
Predicted by Codon_usage
Predicted by Framed"
/codon_start=1
/evidence-not_experimental
/transl_table=11
/product="CONSERVED HYPOTHETICAL PROTEIN"
/protein_id="CAD13534.1"
/db_xref="GI:17427014"
/translation="MRALFTCSRNLRSPTAESIFAHWPNDVTDTSAGLAPDADVPLTA
DQLEWAEVIFVMSRRRLSRQFGLRGRKRVVCLDIPDDYLFPMQPELVALLERRAG
PLYRS"
5188..5721
/genes="RSc0007"
/note="RS01829"

CDS
5188..5721
/genes="RSc0007"
/function="miscellaneous; unknown"
/note="Product confidence : hypothetical
Gene name confidence : hypothetical
Predicted by Framed"
/codon_start=1
/evidence-not_experimental
/transl_table=11
/product="HYPOTHETICAL PROTEIN"
/protein_id="CAD13535.1"
/db_xref="GI:17427015"
/translation="MGRGNDYTWRRKAGEAAEALFYANGWAARAAYALGQRLRVDR
RHVSLVLAELRGLRVPAGMFGVYGRKRPWAVWRGRAFKAREHCYRADRHAARLPD
YAVPPAPRCIRRTAR"
complement(6204..6416)
/genes="rpsU"
/note="RSc0008; RS01830"
complement(6204..6416)
/genes="rpsU"
/function="structural elements; ribosome constituents;
Query Match 2.2%; Score 70.8; DB 1; Length 215050;
Best Local Similarity 51.3%; Pred. No. 3e-05;
Matches 193; Conservative 0; Mismatches 177; Indels 6; Gaps 1;
QY 1255 GGCACGGCGGCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACCTCGGTACC 1314
DB 154091 GCGCCGAGGTGCTGTGGAACGCGCTTCCGTACCTGAGCGCGACCGCGGTG 154150
QY 1315 ACGCTGCTGACAACGGCTCAGGACATCGGTGCTGCGGTGGACAGCAAGTTCCGGTGG 1374
DB 154151 GTGCTGAAGACCACGAGCTGGCTCCGGTACCCTGATGAGGACGCGGAGGCTGG 154210
QY 1375 GGACTGCTGGATGCGGTGAAGCCATGAACGACCGCGCTCTTTCCTTTCGCGGACTTT 1434
DB 154211 GCGCGCTGAACCTGTTTCGGCGCGCGACGCGCTATGCGCGGTTCAACGGCAGCGTGTG 154270
QY 1435 -----ACCGCCGATAGAAAGGTACATCCGATATTCCTACTACTCTTCCGTAAAGACATT 1488
DB 154271 GTGTCGATGATGCTGCTCGGGCGGCTTCAATCCGCGGACACTGGCGCAACGCGGTT 154330
QY 1489 TCAGGACGGCGCGCTGATCAAAAAGGCGGCGGACCAACTGCAACTGCGACGCAACAAC 1548
DB 154331 GCGCGCGGCGCAAGCTTGACGCTCGAGGACCGCGACGCTGCGCTGGCGGCGAGCAAC 154390
QY 1549 ACCTATACGGGCAAAACCATTTATCGAAGGCGGTTTCGCTGCTGTTGTACGGCAACAACAA 1608
DB 154391 ACCTACACGGCGGCGACGAGTTCGCGGCGGCTGCTGCGGCGGCGACTCGGCCACCGCG 154450
QY 1609 TCGGATATGCGCGTCG 1624
DB 154451 TTCGGCAGCGGCGACG 154466

RESULT 37
SMASPPH
LOCUS
DEFINITION
Serratia marcescens DNA for SSP-h1 and SSP-h2, complete cds.
ACCESSION
D78380
VERSION
D78380.1 GI:3688584
KEYWORDS
SSP-h2; SSP-h1.
SOURCE
Serratia marcescens (strain:IFO3046) DNA.
ORGANISM
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Serratia.
REFERENCE
1 (sites)
AUTHORS
Ohnishi,Y., Beppu,T. and Horinouchi,S.
TITLE
Two genes encoding serine protease homologues in Serratia
marcescens and characterization of their products in Escherichia
coli
JOURNAL
J. Biochem. 121 (5), 902-913 (1997)
```

MEDLINE REFERENCE AUTHORS TITLE JOURNAL COMMENT FEATURES	97335937 2 (bases 1 to 7030) Ohnishi, Y. Direct Submission Submitted (24-Nov-1995) Yasuo Ohnishi, The University of Tokyo, 113-8467, Japan (E-mail: ayasuo@hongo.ecc.u-tokyo.ac.jp, Tel: +81-3-3812-2111, Fax: +81-3-3812-0544) On Oct 3, 1998 this sequence version replaced gi:1100762. Sequence updated (24-Sep-1998). Location/Qualifiers 1. .7030 /organism="Serratia marcescens" /strain="IFO3046" /db_xref="taxon:615" 572..3682 /codon_start=1 /transl_table=11 /product="SSP-h1" /protein_id="BAA33455.1" /db_xref="GI:3688585" /translation="MEKKLRGVKSARAAFGWPEIGVHRAOWSAMLICLTATGGAQSS YIENGKADPPASWSSERNAEWGLCAIHADQAYAGYTGKGIKIGFIPQVYAKHPEF AGENKVINLVTEGIREYDPIPVKKGDARFDYDTPSVSDGTGLSGHGHVGGIAAGS RDGAMHGVAFNAQIISAEVDPGPDGIILGNDGAVYQAGWDALVAGSARIINSWG IGITEKDEGGYDPAYPHFTVNDQKQFDQIKQILGTKPGYQCAIDAAARSQVYVTF AAGDGNLNNDPAMAGLAYFVPDIAPNLSVASLQDPTNTGDSYSTSPSRGCVTASL CVSAPCTRVYSIIIGTSVENLTTCYAKYSGTSMAPHVAGSVAVMERFFYLNGAOV AEVLKTTADMGAGVDIGALYGWGMINLNGKAINPGMLVTAEDIEPFRIPTDYGAY PTQVVDLPGVAGDKGDKPTKRVCSVLCGLDFWNSDILSGHGLTKOGIGTGLVLTGN NYSGPTLVNAGRIANGSVTSVAVSQGVGGTSGVSLTARGGTVPAGNSIGTL NVTGDSPEPGRVAVEVPGNSQDRIOSGGAATIGGGEAVTLNENSLNLTQSEVRS LLGOQYTLSDAOGVSGQDAPVAPNVLFLGTLSVQPTGTVLSVGRNCTSPASVAOTP NERVAAAADLAGAPVYESLLASGTAGEARQARQLSGQIHADIASALVNDSKYLR EALNRLKQAGLASSAIIKADEGAWQLLQAMDHSGDANATYQASTTGLVLGLD SAAAADWRLVGTQRTKYSVTLHGSGYKSDSNYHLAAYQKQFALALRGAGYTWHR IDTKRSVYNGQSDRTAKYSARTEQLFAEAGYSVKGEMLNLEFPVNLAYVNFENGI AESGGAALRGDKQHTDASTVTLGLRADTEWQVSPGTTVALRSELGWQHYGGLRGRT GLRFNGGNAPFVVDSPVSRDGMVLKAGAEVAVNENASLSLGYGGLLSQNHQDNSVNA GFTWRF"
BASE COUNT ORIGIN	1370 a 2099 c 2431 g 1130 t
Query Match Best Local Similarity Matches 122; Conservative	2.1%; Score 68; DB 1; Length 7030; 57.5%; Pred No. 0.0001; 90; Mismatches 90; Indels 0; Gaps 0;
QY 1224	CGGAACATCCTTTTCGCCACCCATCTGAACCGGCACGGCGTCTGCTGCTGCAGAAATA 1283
Db 4886	CGGCACCTCGATGGGACGCGCCAGCTCACCGCGCGTGGCGGTGTTGCTGCACGCGTT 4945
QY 1284	CCCGTGGATGACACGACAACTCGTACCAACGCTGCTGACAAACGCTCAGCATCGG 1343
Db 4946	CCCGTACATGACTCGGCGCAGATTGCCAGCTGTGTAACACCCGCCACGACATGGG 5005
QY 1344	TGCAGTCGCGGTGGACAGCAAGTTCGGCTGGGGACTGCTGGATGGGGTAAAGCCATGAA 1403
Db 5006	CGCGCGGCGCATCGACGCGCTCTACGGCTGGGGAATGATCAACCTGGGCAAGGCCATCAA 5065
QY 1404	CGGACCGCGGTCTTTCGTTCCGTCGGCGACTTTTA 1435
Db 5066	CGGCGCGCGCATGTTCCTATACCGTCGAGATA 5097
RESULT 38 AE004006 LOCUS DEFINITION ACCESSION VERSION KEYWORDS SOURCE ORGANISM	11476 bp DNA linear BCT 15-JUN-2001 Xylella fastidiosa 9a5c, section 152 of 229 of the complete genome. AE004006 AE003849 AE004006.1 GI:9106932 Xylella fastidiosa 9a5c. Xylella fastidiosa 9a5c Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group; Xylella. 1 (bases 1 to 11476) Simpson,A.J., Reinach,F.C., Arruda,P., Abreu,F.A., Acencio,M., Alvarenga,R., Alves,L.M., Araya,J.E., Bala,G.S., Baptista,C.S., Barros,M.H., Bonaccorsi,E.D., Bordin,S., Bove,J.M., Briones,M.R., Bueno,M.R., Camargo,A.A., Camargo,L.E., Carraro,D.M., Carrier,H., Colauto,N.B., Colombo,C., Costa,F.F., Costa,M.C., Costa-Neto,C.M., Coutinho,L.L., Cristofani,M., Dias-Neto,E., Docena,C., El-Dorri,H., Facincani,L.P., Ferreira,A.J., Ferreira,V.C., Ferri,J.A., Fraga,J.S., Franca,S.C., Franco,M.C., Frohme,M., Furlan,L.R., Garnier,M., Goldman,G.H., Goldman,M.H., Gomes,S.L., Gruber,A., Ho,P.L., Hohelsel,J.D., Junqueira,M.L., Kemper,E.L., Kitajima,J.P. and Marino,C.L.
REFERENCE AUTHORS	The genome sequence of the plant pathogen Xylella fastidiosa. The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing and Analysis Nature 406 (6792), 151-157 (2000) 20365717 PUBMED 10910347
JOURNAL MEDLINE PUBMED REFERENCE AUTHORS	2 (bases 1 to 11476) Simpson,A.J.G., Reinach,F.C., Arruda,P., Abreu,F.A., Acencio,M., Alvarenga,R., Alves,L.M.C., Araya,J.E., Bala,G.S., Baptista,C.S., Barros,M.H., Bonaccorsi,E.D., Bordin,S., Bove,J.M., Briones,M.R.S., Bueno,M.R.P., Camargo,A.A., Camargo,L.E.A., Carraro,D.M., Carrier,H., Colauto,N.B., Colombo,C., Costa,F.F., Costa,M.C.R., Costa-Neto,C.M., Coutinho,L.L., Cristofani,M., Dias-Neto,E., Docena,C., El-Dorri,H., Facincani,A.P., Ferreira,A.J.S., Ferreira,V.C.A., Ferro,J.A., Fraga,J.S., Franca,S.C., Franco,M.C., Frohme,M., Furlan,L.R., Garnier,M., Goldman,G.H., Goldman,M.H.S., Gomes,S.L., Gruber,A., Ho,P.L., Hohelsel,J.D., Junqueira,M.L., Kemper,E.L., Kitajima,J.P., Krieger,J.E., Kuramae,E.E., Laigret,F., Lambais,M.R., Leite,L.C.C., Lemos,E.G.M., Lemos,M.V.F., Lopes,S.A., Lopes,C.R., Machado,J.A., Machado,M.A., Madeira,A.M.B.N., Madeira,H.M.F., Marino,C.L., Marques,M.V., Martins,E.A.L., Martins,E.M.F., Matsukuma,A.Y., Menck,C.F.M., Miracca,E.C., Miyaki,C.Y., Monteiro-Vitorello,C.B., Moon,D.H., Nagai,M.A., Nascimento,A.L.T.O., Netto,L.E.S., Nhani Jr.,A., Nobrega,F.G., Nunes,L.R., Oliveira,M.A., de Oliveira,M.C., de Oliveira,R.C., Palmieri,D.A., Paris,A., Peixoto,B.R., Pereira,G.A.G., Pereira Jr.,H.A., Pesquero,J.B., Quaggio,R.B., Roberto,P.G., Rodrigues,V., de M. Rosa,A.J., de Rosa Jr.,V.E., de Sa,R.G., Santelli,R.V., Savasaki,H.E., da Silva,A.C.R., da Silva,F.R., da Silva,A.M., Silva Jr.,W.A., da Silveira,J.F., Silvestri,M.L.Z., Siqueira,W.J., de Souza,A.A., de Souza,A.P., Terenzi,M.F., Truffi,D., Tsai,S.M., Tshako,M.H., Vallada,H., Van Sluys,M.A., Verjovski-Almeida,S., Vettore,A.L., Zago,M.A., Zatz,M., Meidanis,J. and Setubal,J.C. Direct Submission





Query Match 2.1%; Score 67; DB 1; Length 11476;  
Best Local Similarity 55.9%; Pred. No. 0.00018;  
Matches 127; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 1224 CGGAACATCTTTTCGCGACCACATCGTAACCGGCAGCGCGCTCTGCTGTGCAGAATA 1283  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 4978 CGCAGCTCCTTTTTCAGCCGCCCTTGATCTGGTGCACGCCGCTTAGTGTGGAGGTATT 5037  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1284 CCGGTGGATGAGCAACAACCTCGTACCCAGCTGTGCAGACGGCTCAGGACATCGG 1343  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 5038 TCCATACTTTAATAACGATAGTACGTCGTCAGACATTCGTTGGACGCCACCATCTGG 5097  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1344 TGCAGTCGCGGTGGACAGCAAGTTCGCTGGGGACTGCTGTGATCGGGTAAAGCCATGAA 1403  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 5098 TGCTCTCGCGTGGACGAAGTAGTATCGGCTATCGGCTCGAATGTGGCAAGCGGTGCG 5157  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1404 CGGACCGCGCTCTTTTCGTTCCGGCACTTTACCGCCCATACGAAG 1450  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 5158 TGGCCCTGCCAAGTTTGATGGGGCGATGTCAAGGTCAACTTCAATG 5204  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 39  
AF216702 5423 bp DNA linear BCT 06-MAR-2001  
LOCUS Pseudomonas fluorescens PrtB (prtB) and lipase (lipA) genes,  
complete cds.  
DEFINITION AF216702  
VERSION AF216702.1 GI:8895499  
KEYWORDS Pseudomonas fluorescens.  
SOURCE Pseudomonas fluorescens.  
ORGANISM Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
Pseudomonas.

REFERENCE 1 (bases 1 to 5423)  
Woods,R.G., Burger,M., Beven,C.A. and Beacham,I.R.  
The aprX-lipA operon of Pseudomonas fluorescens B52: a molecular  
analysis of metalloprotease and lipase production  
Microbiology 147 (Pt 2), 345-354 (2001)

JOURNAL 21097273  
MEDLINE  
PUBMED 11158351

REFERENCE 2 (bases 1 to 5423)  
Woods,R.G., Burger,M., Beven,C.-A. and Beacham,I.R.  
Direct Submission  
Submitted (16-DEC-1999) Health Science, Griffith University,  
Parklands Drive, Goldcoast, Qld 4215, Australia  
Location/Qualifiers  
FEATURES Source  
1..5423  
/organism="Pseudomonas fluorescens"  
/strain="B52"  
/db\_xref="taxon:294"  
575..3685  
/gene="prtB"  
575..3685  
/gene="prtB"  
/note="similar to serine protease"  
/codon\_start=1  
/transl\_table=11  
/product="prtB"  
/protein\_id="AAF80995.1"  
/db\_xref="GI:8895500"

gene  
CDS

AVGERLNGEMGATSETLDRGNVWKALCAWGKTDSDRSTAGTYTSLIGMLVGVDG  
LDEATRILGVAGYSDTLSNMGNDTHSRVSFYHFGAYAGHEIFLARLISGGATYSWH  
ADVRDLQYDGVAGKVQKVDASSTQVFGEAAYRVNLQPLAEFPFLAFVHLDTGDF  
TEKGDAALKGHDDRDVLSTLCMRALKTFVNVDHQOLEVSGTCLGWHLNSGTSEQ  
HLFASGPSPEFAVESAPMVRDAALVGARVSLALSKEARYNFYDINGLLASKDKVHGIVG  
SLDWAF"  
3814..5244  
/gene="lipA"  
3814..5244  
/gene="lipA"  
/codon\_start=1  
/transl\_table=11  
/product="lipase"  
/protein\_id="AAF80996.1"  
/db\_xref="GI:8895501"  
/translation="MGIFYDNKLGTGSKTLFDAMAAILTVSYHNLDNGPAVGYOHNG  
LGGLPATLVGALLGSQTDSOGVIPGPWPNDPSEKALEAVQAGWTPLISASALGYAGR  
VDARGTFEGKAGYTITAQVEVLGRIDDAKKLEIGIGFRGTSPRETLISDSIGDLIS  
DLAALGPKDYKNYAGEAFGGLLKNVAGAHGLTGKDVVVSHSLGGLAVNSMD  
LSNKWAGFYKDANYAVAYSQTSKADKVLNIENYENDPVFRALDGSSFLSLGSHVK  
PHESTTDNISFNHDHYASTLWNLPFSIVNLTPTWSHLPTAYGDGMTRLSEGFYDM  
TRDSTIVIANLSDPARANTWODLNRAEPHKNTFTIGSDGNDLIQGNGADFIBGG  
KGNDFIRDNSCHNTELEPSGHFNDRVLYGYQPTDKLFKPDQVSGSTDLRDHAKVVGADTV  
LTPGADSVTLVGVGHGGLWTEGVIG"

BASE COUNT 1093 a 1659 c 1734 g 937 t  
ORIGIN

Query Match 2.0%; Score 65; DB 1; Length 5423;  
Best Local Similarity 56.2%; Pred. No. 0.0005;  
Matches 122; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 1225 GGAACATCCTTTTTTCGCGCACCCATCGTAACCGGCAGCGCGCTCTGCTGTGCAGAAATAC 1284  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 1679 GCACCTTCATGCGCGCACCCACGTCGCGCGCGCTGCGCGGTGTTGATGGAACGCTTC 1738  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1285 CCGTGGATGAGCAACGACACCTGCGTACCAGCTGTGTGACACGGCTCAGGACATCGGT 1344  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 1739 CCGTACATGACGCGCATCAGTCTCCACCGTGTCAAACACCGCCGACCGAACCTCCGC 1798  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1345 CGAGTCGCGGTGCACGCAAGTTCGCTGGGACTGCTGGATGCGGGTAAAGCCATGAAC 1404  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 1799 GGCGCAGGATGACGCGCTGTACCGCTGGGCGATGATCAACCTGGGCAAGCCGCGCAAC 1858  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1405 GGACCGCGCTCCTTTTCGCTGCGGCATCTTACCGCG 1441  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 1859 GGCCCGGAATGTTTCATTACCGCGCTGATATCCCG 1895  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 40  
E32986  
LOCUS E32986 10732 bp DNA linear PAT 18-JUN-2001  
DEFINITION Gene encoding cellulose synthesizer.  
ACCESSION E32986  
VERSION E32986.1 GI:13022340  
KEYWORDS JP 2000060568-A/1.  
SOURCE Vigna angularis.  
ORGANISM Vigna angularis  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Vigna.  
REFERENCE 1 (bases 1 to 10732)  
Koichi,M., Tomohiko,K., Shigeru,S. and Daisuke,S.  
AUTHORS Gene encoding cellulose synthesizer  
TITLE Patent: JP 2000060568-A 1 29-FEB-2000;  
JOURNAL KOICHI MIZUNO, MITSUI GIYOUSAI SHOKUBUTSU BIO KENKYUSHO  
COMMENT OS Vigna angularis  
PN JP 2000060568-A/1  
PD 29-FEB-2000  
PF 26-AUG-1998 JP 1998239998  
PR  
PI KOICHI MIZUNO, TOMOHIKO KATO, SHIGERU SATO, DAISUKE SHIBATA PC  
C12N15/09,A01H5/00,C12N9/10,C12N9/10,C12P21/02// PC

Search completed: January 27, 2003, 19:00:20  
Job time : 7910 secs

**THIS PAGE BLANK (USPTO)**

09/830433

FILE 'HCAPLUS, MEDLINE, BIOSIS, EMBASE, WPIDS, CONFSCI, SCISEARCH,  
JICST-EPLUS, JAPIO' ENTERED AT 10:13:04 ON 28 JAN 2003

- Author(s)

L1 105 SEA ABB=ON PLU=ON "AUJAME L"?/AU  
L2 29 SEA ABB=ON PLU=ON "BOUCHARDON A"?/AU  
L3 79 SEA ABB=ON PLU=ON ("RENAULD MONGENIE G"? OR "MONGENIE  
RENAULD G"? OR "MONGENIE G"? OR "RENAULD G"?)/AU  
L4 65 SEA ABB=ON PLU=ON "ROKBI B"?/AU  
L5 309 SEA ABB=ON PLU=ON "NASSIF X"?/AU  
L6 208 SEA ABB=ON PLU=ON "TINSLEY C"?/AU  
L7 1372 SEA ABB=ON PLU=ON "PERRIN A"?/AU  
L8 2 SEA ABB=ON PLU=ON L1 AND L2 AND L3 AND L4 AND L5 AND  
L6 AND L7  
L9 5 SEA ABB=ON PLU=ON L1 AND (L2 OR L3 OR L4 OR L5 OR L6  
OR L7)  
L10 2 SEA ABB=ON PLU=ON L2 AND (L3 OR L4 OR L5 OR L6 OR L7)  
L11 8 SEA ABB=ON PLU=ON L3 AND (L4 OR L5 OR L6 OR L7)  
L12 2 SEA ABB=ON PLU=ON L4 AND (L5 OR L6 OR L7)  
L13 52 SEA ABB=ON PLU=ON L5 AND (L6 OR L7)  
L14 19 SEA ABB=ON PLU=ON L6 AND L7  
L15 2079 SEA ABB=ON PLU=ON L1 OR L2 OR L3 OR L4 OR L5 OR L6 OR  
L7  
L16 293 SEA ABB=ON PLU=ON (L13 OR L15) AND NEISSER?  
L17 66 SEA ABB=ON PLU=ON L16 AND (VACCIN? OR IMMUNIS? OR  
IMMUNIZ?)  
L18 26 SEA ABB=ON PLU=ON L17 AND PATHOGEN?  
L19 52 SEA ABB=ON PLU=ON L8 OR L9 OR L10 OR L11 OR L12 OR L14  
OR L18  
L20 25 DUP REM L19 (27 DUPLICATES REMOVED)

L20 ANSWER 1 OF 25 HCAPLUS COPYRIGHT 2003 ACS DUPLICATE 1  
ACCESSION NUMBER: 2002:908217 HCAPLUS  
TITLE: Comparative genomics identifies the genetic  
islands that distinguish *Neisseria meningitidis*,  
the agent of cerebrospinal meningitis, from  
other *Neisseria* species  
AUTHOR(S): Perrin, Agnes; Bonacorsi, Stephane;  
Carbonnelle, Etienne; Talibi, Driss; Dessen,  
Philippe; Nassif, Xavier; Tinsley, Colin  
CORPORATE SOURCE: INSERM U5701 Faculte de Medecine Necker, Paris,  
75015, Fr.  
SOURCE: Infection and Immunity (2002), 70(12), 7063-7072  
CODEN: INFIBR; ISSN: 0019-9567  
PUBLISHER: American Society for Microbiology  
DOCUMENT TYPE: Journal  
LANGUAGE: English

AB *Neisseria meningitidis* colonizes the nasopharynx and, unlike  
commensal *Neisseria* species, is capable of entering the bloodstream,  
crossing the blood-brain barrier, and invading the meninges. The  
other pathogenic *Neisseria* species, *Neisseria gonorrhoeae*, generally  
causes an infection which is localized to the genitourinary tract.  
In order to investigate the genetic basis of this difference in  
disease profiles, we used a strategy of genomic comparison. We used  
DNA arrays to compare the genome of *N. meningitidis* with those of *N.*  
*gonorrhoeae* and *Neisseria lactamica*, a commensal of the nasopharynx.  
We thus identified sequences conserved among a representative set of  
virulent strains which are either specific to *N. meningitidis* or  
shared with *N. gonorrhoeae* but absent from *N. lactamica*. Though

**THIS PAGE BLANK (USPTO)**

09/830433

these bacteria express dramatically different pathogenicities, these meningococcal sequences were limited and, in contrast to what has been found in other pathogenic bacterial species, they are not organized in large chromosomal islands.

REFERENCE COUNT: 39 THERE ARE 39 CITED REFERENCES AVAILABLE  
FOR THIS RECORD. ALL CITATIONS AVAILABLE  
IN THE RE FORMAT

L20 ANSWER 2 OF 25 MEDLINE  
ACCESSION NUMBER: 2002158631 MEDLINE  
DOCUMENT NUMBER: 21887396 PubMed ID: 11890539  
TITLE: Genomics of **Neisseria meningitidis**.  
AUTHOR: **Nassif Xavier**  
CORPORATE SOURCE: INSERM U411, Faculte de Medecine Necker-Enfants  
Malades, Universite Rene Descartes, Paris, France..  
nassif@necker.fr  
SOURCE: INTERNATIONAL JOURNAL OF MEDICAL MICROBIOLOGY, (2002  
Feb) 291 (6-7) 419-23. Ref: 28  
Journal code: 100898849. ISSN: 1438-4221.  
PUB. COUNTRY: Germany: Germany, Federal Republic of  
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)  
General Review; (REVIEW)  
(REVIEW, TUTORIAL)  
LANGUAGE: English  
FILE SEGMENT: Priority Journals  
ENTRY MONTH: 200209  
ENTRY DATE: Entered STN: 20020314  
Last Updated on STN: 20020928  
Entered Medline: 20020927

AB An important feature of disease caused by **Neisseria**  
meningitidis is the propensity to invade the meninges. Much progress  
has been made in our understanding of how this **pathogen**  
circumvents the physical properties of this cellular barrier. This  
review will address the new possibilities offered by the recent  
availability of meningococcal genome sequences.

L20 ANSWER 3 OF 25 MEDLINE DUPLICATE 2  
ACCESSION NUMBER: 2002712859 MEDLINE  
DOCUMENT NUMBER: 22362872 PubMed ID: 12474388  
TITLE: **Neisseria** microarrays.  
AUTHOR: **Tinsley Colin R; Perrin Agnes;**  
Borezee Elise; Nassif Xavier  
CORPORATE SOURCE: INSERM U570, Faculte de Medecine Necker-Enfants  
Malades, 75730 Paris, France.  
SOURCE: METHODS IN ENZYMOLOGY, (2002) 358 188-207.  
Journal code: 0212271. ISSN: 0076-6879.  
PUB. COUNTRY: United States  
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)  
LANGUAGE: English  
FILE SEGMENT: Priority Journals  
ENTRY MONTH: 200301  
ENTRY DATE: Entered STN: 20021217  
Last Updated on STN: 20030122  
Entered Medline: 20030121

L20 ANSWER 4 OF 25 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.  
ACCESSION NUMBER: 2001:477334 BIOSIS  
DOCUMENT NUMBER: PREV200100477334

**THIS PAGE BLANK (USPTO)**



09/830433

TITLE: Hemoglobin receptors from **neisseriae**.  
AUTHOR(S): Stojiljkovic, Igor; So, Magdalene; Hwa, Vivian;  
Heffron, Fred; Nassif, Xavier (1)  
CORPORATE SOURCE: (1) Paris France  
ASSIGNEE: Oregon Health Sciences University  
PATENT INFORMATION: US 6277382 August 21, 2001  
SOURCE: Official Gazette of the United States Patent and  
Trademark Office Patents, (Aug. 21, 2001) Vol. 1249,  
No. 3, pp. No Pagination. e-file.  
ISSN: 0098-1133.  
DOCUMENT TYPE: Patent  
LANGUAGE: English  
AB The present invention relates to novel bacterial hemoglobin receptor  
proteins and genes that encode such proteins. The invention is  
directed toward the isolation, characterization, diagnostic and  
therapeutic use of bacterial hemoglobin receptor proteins, nucleic  
acids encoding such proteins, recombinant expression constructs  
comprising such nucleic acids and cells transformed therewith, and  
antibodies and epitopes of such hemoglobin receptor proteins. The  
invention relates particularly to hemoglobin receptor proteins and  
genes encoding such proteins from **Neisseria** species,  
especially *N. meningitidis* and serotypes thereof, and *N.*  
*gonorrhoeae*. Methods for the diagnostic and therapeutic use of the  
proteins, epitopes, antibodies and nucleic acids of the invention  
are also provided, including the use of proteins, epitopes,  
antibodies and nucleic acids of the invention for the production of  
**vaccines** effective in providing **immunization** of  
human against infection by **pathogenic** bacteria of  
**Neisseria** species.

L20 ANSWER 5 OF 25 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.

ACCESSION NUMBER: 2001:334753 BIOSIS  
DOCUMENT NUMBER: PREV200100334753  
TITLE: Preparation of a multicombinatorial library of  
antibody gene expression vectors.  
AUTHOR(S): Sodoyer, Regis (1); Aujame, Luc; Geoffroy,  
Frederique; Bouchardon, Annabelle  
CORPORATE SOURCE: (1) Saint Foy les Lyon France  
ASSIGNEE: Pasteur Merieux Serums & Vaccins, Lyon  
Cedex, France  
PATENT INFORMATION: US 6174708 January 16, 2001  
SOURCE: Official Gazette of the United States Patent and  
Trademark Office Patents, (Jan. 16, 2001) Vol. 1242,  
No. 3, pp. No Pagination. e-file.  
ISSN: 0098-1133.  
DOCUMENT TYPE: Patent  
LANGUAGE: English  
AB On the basis of a first repertoire of genes coding for a population  
of one of two kinds of polypeptides capable of being optionally  
covalently combined, particularly variable regions of either the  
antibody light chain type or the antibody heavy chain type, and at  
least one gene coding for the other type of polypeptide,  
particularly a variable region of the other type, an antibody chain  
or preferably a second repertoire of genes coding for a population  
of said other type, the genes from the first repertoire are inserted  
into a first vector to form a population of vectors carrying the  
various genes of said first repertoire, and said gene of said other  
type or the genes from said second repertoire is/are inserted into a

**THIS PAGE BLANK (USPTO)**

09/830433

second vector. Both starting vectors have means enabling each to exchange one part by one or more irreversible recombinations to generate recombinant final vectors of which one contains a gene of one of said types and a gene of the other type.

L20 ANSWER 6 OF 25 WPIDS (C) 2003 THOMSON DERWENT  
ACCESSION NUMBER: 2001-082916 [10] WPIDS  
DOC. NO. NON-CPI: N2001-063334  
DOC. NO. CPI: C2001-024200  
TITLE: Immunogenic polypeptides derived from  
**Neisseria meningitidis** and the nucleic  
acids that encode them, useful for diagnosing and  
**vaccinating** against **Neisseria**  
infections e.g. bacteremia and meningitis.  
DERWENT CLASS: B04 D16 S03  
INVENTOR(S): **NASSIF, X; TINSLEY, C; ACHTMAN,**  
M; KLEE, S; MERKER, P  
PATENT ASSIGNEE(S): (INRM) INSERM INST NAT SANTE & RECH MEDICALE;  
(PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN  
COUNTRY COUNT: 95  
PATENT INFORMATION:

PATENT NO	KIND	DATE	WEEK	LA	PG
EP 1069133	A1	20010117	(200110)*	EN	232
R: AL AT BE CH CY DE DK ES FI FR GB GR IE IT LI LT LU LV MC MK NL PT RO SE SI					
WO 2001004150	A2	20010118	(200110)	EN	
RW: AT BE CH CY DE DK EA ES FI FR GB GH GM GR IE IT KE LS LU MC MW MZ NL OA PT SD SE SL SZ TZ UG ZW					
W: AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CR CU CZ DE DK DM DZ EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR LS LT LU LV MA MD MG MK MN MW MX MZ NO NZ PL PT RO RU SD SE SG SI SK SL TJ TM TR TT TZ UA UG US UZ VN YU ZA ZW					
AU 2000068254	A	20010130	(200127)		
EP 1194446	A2	20020410	(200232)	EN	
R: AL AT BE CH CY DE DK ES FI FR GB GR IE IT LI LT LU LV MC MK NL PT RO SE SI					

APPLICATION DETAILS:

PATENT NO	KIND	APPLICATION	DATE
EP 1069133	A1	EP 1999-401764	19990713
WO 2001004150	A2	WO 2000-EP6943	20000705
AU 2000068254	A	AU 2000-68254	20000705
EP 1194446	A2	EP 2000-956222	20000705
		WO 2000-EP6943	20000705

FILING DETAILS:

PATENT NO	KIND	PATENT NO
AU 2000068254	A Based on	WO 200104150
EP 1194446	A2 Based on	WO 200104150

PRIORITY APPLN. INFO: EP 1999-401764 19990713

Searcher : Shears 308-4994

**THIS PAGE BLANK (USPTO)**

AN 2001-082916 [10] WPIDS

AB EP 1069133 A UPAB: 20010220

NOVELTY - Immunologically active polypeptides (I) derived from the Gram negative bacteria *Neisseria meningitidis*, and the nucleic acids (II) that encode them, are new.

DETAILED DESCRIPTION - INDEPENDENT CLAIMS are included for the following:

- (1) an isolated polypeptide (I) comprising an amino acid sequence that has at least 70% identity to 44 defined amino acid sequences ((A1)-(A44)) given in the specification;
- (2) an immunogenic fragment of (I) which comprises (A1)-(A44);
- (3) an isolated polynucleotide (II) comprising a nucleotide sequence encoding (I) (which has at least 70% to (A1)-(A44) over its entire length), or a sequence complementary to (II);
- (4) an expression vector (III) or a recombinant live microorganism comprising (II);
- (5) a host cell (IV) comprising (III), or a membrane of (IV), that expresses a polypeptide comprising an amino acid sequence with at least 70% identity to (A1)-(A44);
- (6) a process (V) for producing a polypeptide comprising an amino acid sequence with at least 70% identity to (A1)-(A44), comprising culturing the host cell (IV) under suitable conditions for expression of the polypeptide and recovering the polypeptide from the culture medium;
- (7) a process (VI) for expressing the polynucleotide (II), comprising transforming a host cell with an expression vector comprising (II) and culturing the host cell under conditions suitable for expression of the polypeptide;
- (8) **vaccine** compositions (VII) comprising (I) and/or (II);
- (9) antibody (VIII) immuno-specific for (I); and
- (10) a method for diagnosing a *Neisseria* infection, comprising identifying (I) or (VIII) in a sample from the subject animal.

ACTIVITY - Antibacterial.

MECHANISM OF ACTION - **Vaccine**.

Rabbit antiserum produced in response to **vaccination** with the polypeptides killed 65% of parenterally administered meningococcus (strain 8013) with in 20 minutes of contact and all of the bacteria within 60 minutes. Pre-immune serum (taken prior to **immunization**) was found to have killed no bacteria after 20 minutes and only half after 60 minutes.

USE - The nucleic acids and the polypeptides they encode may be used to **vaccinate** subjects against infection by *Neisseria meningitidis* bacteria according to standard methodologies. The antibodies produced in response to the polypeptides and/or polynucleotides may also be used to treat *N. meningitidis* infections or as diagnostic reagents in immunoassays to detect infections (claimed). *N. meningitidis* is a **pathogen** involved in, for example, bacteremia and meningitis.

Dwg.0/50

L20 ANSWER 7 OF 25 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.

ACCESSION NUMBER: 2001:443973 BIOSIS

DOCUMENT NUMBER: PREV200100443973

TITLE: Production of *Neisseria meningitidis* transferrin-binding protein B by recombinant *Bordetella pertussis*.

**THIS PAGE BLANK (USPTO)**

09/830433

AUTHOR(S): Coppens, Isabelle; Alonso, Sylvie; Antoine, Rudy;  
Jacob-Dubuisson, Francoise; **Renauld-Mongenie,**  
**Genevieve**; Jacobs, Eric; Locht, Camille (1)  
CORPORATE SOURCE: (1) Laboratoire de Microbiologie Genetique et  
Moleculaire, INSERM U447, Institut Pasteur de Lille,  
1 Rue du Prof. Calmette, F-59019, Lille Cedex:  
camille.locht@pasteur-lille.fr France  
SOURCE: Infection and Immunity, (September, 2001) Vol. 69,  
No. 9, pp. 5440-5446. print.  
ISSN: 0019-9567.  
DOCUMENT TYPE: Article  
LANGUAGE: English  
SUMMARY LANGUAGE: English

AB **Neisseria** meningitidis serogroup B infections are among  
the major causes of fulminant septicemia and meningitis, especially  
severe in young children, and no broad **vaccine** is  
available yet. Because of poor immunogenicity of the serogroup B  
capsule, many efforts are now devoted to the identification of  
protective protein antigens. Among those are PorA and, more  
recently, transferrin-binding protein B (TbpB). In this study, TbpB  
of *N. meningitidis* was genetically fused to the N-terminal domain of  
the *Bordetella pertussis* filamentous hemagglutinin (FHA), and the  
fha-tbpB hybrid gene was expressed in *B. pertussis* either as a  
plasmid-borne gene or as a single copy inserted into the chromosome.  
The hybrid protein was efficiently secreted by the recombinant  
strains, despite its large size, and was recognized by both anti-FHA  
and anti-TbpB antibodies. A single intranasal administration of  
recombinant virulent or pertussis-toxin-deficient, attenuated *B.*  
*pertussis* to mice resulted in the production of antigen-specific  
systemic immunoglobulin G (IgG), as well as local IgG and IgA. The  
anti-TbpB serum antibodies were of the IgG1, IgG2a, and IgG2b  
isotypes and were found to express complement-mediated bactericidal  
activity against *N. meningitidis*. These observations indicate that  
recombinant *B. pertussis* may be a promising vector for the  
development of a mucosal **vaccine** against serogroup B  
meningococci.

L20 ANSWER 8 OF 25 MEDLINE DUPLICATE 3  
ACCESSION NUMBER: 2001232892 MEDLINE  
DOCUMENT NUMBER: 21112216 PubMed ID: 11173033  
TITLE: Meningococcal **pathogenesis**: at the boundary  
between the pre- and post-genomic eras.  
AUTHOR: **Tinsley C; Nassif X**  
CORPORATE SOURCE: INSERM U411, Faculte de Medecine Necker - Enfants  
Malades, Universite Rene Descartes, 75730 Cedex 15,  
Paris, France.  
SOURCE: CURRENT OPINION IN MICROBIOLOGY, (2001 Feb) 4 (1)  
47-52. Ref: 49  
Journal code: 9815056. ISSN: 1369-5274.  
PUB. COUNTRY: England: United Kingdom  
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)  
General Review; (REVIEW)  
(REVIEW, TUTORIAL)  
LANGUAGE: English  
FILE SEGMENT: Priority Journals  
ENTRY MONTH: 200105  
ENTRY DATE: Entered STN: 20010517  
Last Updated on STN: 20010517

Searcher : Shears 308-4994

**THIS PAGE BLANK (USPTO)**



09/830433

Entered Medline: 20010503

AB Meningococcal disease remains an important public health burden worldwide and, indeed, cause of death, particularly in poorer countries. The rapidly progressive nature of infections means that antibiotic therapy often comes too late. **Vaccines** are of limited efficacy in infants, one of the most vulnerable age groups, and do not exist for bacteria of serogroup B. Hence, much remains to be achieved in terms of **vaccine** design and the understanding of the **pathogenesis** of meningococcal disease. The causative bacterium, **Neisseria meningitidis**, is usually a commensal of the nasopharynx. Factors that lead to the invasion of the bloodstream, often followed by the crossing of the blood-brain barrier and meningitis, may be partly host- and partly bacterium-dependent, but are ill-understood. It is hoped that, taken together with the fundamental knowledge gained from biochemical and genetic studies, the huge amount of new information made available with the recent publication of the genome sequences will help to unlock more of the secrets of the lifestyle and **pathogenic** potential of this still poorly understood **pathogen**.

L20 ANSWER 9 OF 25 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.

ACCESSION NUMBER: 2002:211711 BIOSIS

DOCUMENT NUMBER: PREV200200211711

TITLE: New pathogenicity factors from genomic comparison.

AUTHOR(S): **Tinsley, C. R. (1); Perrin, A. (1)**  
; Nassif, X. (1)

CORPORATE SOURCE: (1) Faculte de Medecine Necker, INSERM U411, 156 Rue de Vaugirard, 75015, Paris France

SOURCE: International Journal of STD & AIDS, (2001) Vol. 12, No. Supplement 2, pp. 22-23. print.  
Meeting Info.: International Congress of Sexually Transmitted Infections Berlin, Germany June 24-27, 2001 International Union Against Sexually Transmitted Infections  
. ISSN: 0956-4624.

DOCUMENT TYPE: Conference

LANGUAGE: English

L20 ANSWER 10 OF 25 MEDLINE

ACCESSION NUMBER: 2000210556 MEDLINE

DOCUMENT NUMBER: 20210556 PubMed ID: 10755929

TITLE: Microbiology. A furtive **pathogen** revealed.

COMMENT: Comment on: Science. 2000 Mar 10;287(5459):1809-15

Comment on: Science. 2000 Mar 10;287(5459):1816-24

AUTHOR: **Nassif X**

CORPORATE SOURCE: INSERM Unit 411, Facult| de M|decine Necker-Enfants Malades, 156 Rue de Vaugirard, 75015 Paris, France..  
nassif@necker.fr

SOURCE: SCIENCE, (2000 Mar 10) 287 (5459) 1767-8.  
Journal code: 0404511. ISSN: 0036-8075.

PUB. COUNTRY: United States

DOCUMENT TYPE: Commentary  
Journal; Article; (JOURNAL ARTICLE)

LANGUAGE: English

FILE SEGMENT: Priority Journals

ENTRY MONTH: 200004

ENTRY DATE: Entered STN: 20000413

Last Updated on STN: 20000413

Searcher : Shears 308-4994

**THIS PAGE BLANK (USPTO)**

09/830433

Entered Medline: 20000403

L20 ANSWER 11 OF 25 HCAPLUS COPYRIGHT 2003 ACS DUPLICATE 4  
ACCESSION NUMBER: 2000:314839 HCAPLUS  
DOCUMENT NUMBER: 132:330635  
TITLE: Genes and proteins specific for Neisseria  
meningitidis and their use in vaccination  
INVENTOR(S): **Aujame, Luc; Bouchardon,  
Annabelle; Renauld-Mongenie,  
Genevieve; Rokbi, Bachra;  
Nassif, Xavier; Tinsley, Colin  
; Perrin, Agnes**  
PATENT ASSIGNEE(S): Pasteur Merieux Serums et Vaccins, Fr.; Institut  
National de la Sante et de la Recherche Medicale  
(INSERM)  
SOURCE: PCT Int. Appl., 187 pp.  
CODEN: PIXXD2  
DOCUMENT TYPE: Patent  
LANGUAGE: French  
FAMILY ACC. NUM. COUNT: 1  
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2000026375	A2	20000511	WO 1999-FR2643	19991028
WO 2000026375	A3	20000817		
W:	AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM			
RW:	GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG			
FR 2785293	A1	20000505	FR 1998-13693	19981030
FR 2785293	B1	20020705		
AU 9963479	A1	20000522	AU 1999-63479	19991028
EP 1129195	A2	20010905	EP 1999-950875	19991028
R:	AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, SI, LT, LV, FI, RO			

PRIORITY APPLN. INFO.: FR 1998-13693 A 19981030  
WO 1999-FR2643 W 19991028

AB The invention concerns nucleic acids coding for polypeptides specific for Neisseria meningitidis, the corresponding polypeptides, and their diagnostic and therapeutic applications. Thus, genes and proteins found in N. meningitidis but not in N. lactamica were identified and sequenced.

L20 ANSWER 12 OF 25 HCAPLUS COPYRIGHT 2003 ACS DUPLICATE 5  
ACCESSION NUMBER: 2000:658423 HCAPLUS  
DOCUMENT NUMBER: 133:248066  
TITLE: Sequence and recombinant production of  
**Neisseria meningitidis** serotype B gene  
hmbR hemoglobin receptor  
INVENTOR(S): **Stojiljkovic, Igor; So, Magdalene; Hwa, Vivian;  
Heffron, Fred; Nassif, Xavier**  
PATENT ASSIGNEE(S): USA

Searcher : Shears 308-4994

**THIS PAGE BLANK (USPTO)**

09/830433

SOURCE: U.S., 87 pp., Cont.-in-part of U.S. 5,698,438.  
 CODEN: USXXAM  
 DOCUMENT TYPE: Patent  
 LANGUAGE: English  
 FAMILY ACC. NUM. COUNT: 2  
 PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 6121037	A	20000919	US 1995-537361	19951002
US 5698438	A	19971216	US 1994-326670	19941018
CA 2203116	AA	19960425	CA 1995-2203116	19951017
WO 9612020	A2	19960425	WO 1995-US13623	19951017
WO 9612020	A3	19960523		
W: AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LT, LU, LV, MD, MG, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TT				
RW: KE, MW, SD, SZ, UG, AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG				
HU 77048	A2	19980302	HU 1997-1963	19951017
JP 10508469	T2	19980825	JP 1995-513482	19951017
NO 9701768	A	19970603	NO 1997-1768	19970417
FI 9701634	A	19970616	FI 1997-1634	19970417
US 6123942	A	20000926	US 1997-990470	19971215

PRIORITY APPLN. INFO.:  
 US 1994-326670 A2 19941018  
 US 1995-537361 A 19951002  
 WO 1995-US13623 W 19951017

AB The invention claims a DNA mol. encoding a Hb receptor protein from **Neisseria meningitidis** serotype B. The invention also claims the use of said DNA in construction of an expression construct and cells transformed with said expression construct for the recombinant prodn. of Hb receptor. The invention further provides a nucleic acid mol. fully complementary to N. meningitis serotype B Hb receptor DNA as a diagnostic agent. The DNA sequence, as well as the corresponding amino acid sequence of N. meningitis serotype B Hb receptor are provided. In the example section, the invention provided methods used clone Hb receptor genes (hmbR) from N. meningitis serotypes A and C, as well as from N. gonorrhoeae. The example also provided the DNA sequences, as as corresponding amino acid sequences, of these cloned Hb receptor. The invention further showed that: (1) the hmbR gene product is a Hb receptor in N. meningitis; (2) N. meningitis Hb receptor of the invention is TonB-dependent and (3) an IS1106 element is located downstream of hmbR gene in N. meningitis serotype C. Still further the invention discussed use of the proteins, epitopes, antibodies and nucleic acids of the invention for the prodn. of **vaccines** effective in providing **immunization** of a human against infection by **pathogenic** bacteria of **Neisseria** species.

REFERENCE COUNT: 49 THERE ARE 49 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L20 ANSWER 13 OF 25 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.  
 ACCESSION NUMBER: 2001:231857 BIOSIS  
 DOCUMENT NUMBER: PREV200100231857

Searcher : Shears 308-4994

**THIS PAGE BLANK (USPTO)**

09/830433

TITLE: Bacterial hemoglobin receptor gene.  
AUTHOR(S): Stojiljkovic, Igor (1); So, Magdalene; Hwa, Vivian; Heffron, Fred; **Nassif, Xavier**  
CORPORATE SOURCE: (1) Portland, OR USA  
ASSIGNEE: Oregon Health Sciences University  
PATENT INFORMATION: US 6123942 September 26, 2000  
SOURCE: Official Gazette of the United States Patent and Trademark Office Patents, (Sep. 26, 2000) Vol. 1238, No. 4, pp. No Pagination. e-file.  
ISSN: 0098-1133.  
DOCUMENT TYPE: Patent  
LANGUAGE: English  
AB The present invention relates to a novel bacterial hemoglobin receptor protein and genes that encode such a protein. The invention is directed toward the isolation, characterization, diagnostic and therapeutic use of a bacterial hemoglobin receptor protein, nucleic acid encoding such a protein, recombinant expression constructs comprising such nucleic acids and cells transformed therewith, and antibodies and epitopes of such hemoglobin receptor proteins. The invention relates particularly to hemoglobin receptor proteins and genes encoding such proteins from **Neisseria** species, especially *N. meningitidis*. Methods for the diagnostic and therapeutic use of the proteins, epitopes, antibodies and nucleic acids of the invention are also provided, including the use of the proteins, epitopes, antibodies and nucleic acids of the invention for the production of **vaccines** effective in providing **immunization** of a human against infection by **pathogenic bacteria** of **Neisseria** species.

L20 ANSWER 14 OF 25 HCAPLUS COPYRIGHT 2003 ACS DUPLICATE 6  
ACCESSION NUMBER: 2000:603693 HCAPLUS  
DOCUMENT NUMBER: 134:52089  
TITLE: Allelic diversity of the two transferrin binding protein B gene isotypes among a collection of *Neisseria meningitidis* strains representative of serogroup B disease: implication for the composition of a recombinant TbpB-based vaccine  
AUTHOR(S): **Rokbi, Bachra; Renauld-Mongenie, Genevieve**; Mignon, Michele; Danve, B.; Poncet, David; Chabanel, Christophe; Caugant, Dominique A.; Quentin-Millet, Marie-Jose  
CORPORATE SOURCE: Aventis Pasteur, Marcy-L'Etoile, 69280, Fr.  
SOURCE: Infection and Immunity (2000), 68(9), 4938-4947  
CODEN: INFIBR; ISSN: 0019-9567  
PUBLISHER: American Society for Microbiology  
DOCUMENT TYPE: Journal  
LANGUAGE: English  
AB The distribution of the two isotypes of *tbpB* in a collection of 108 serogroup B meningococcal strains belonging to the four major clonal groups assocd. with epidemic and hyperendemic disease (the ET-37 complex, the ET-5 complex, lineage III, and cluster A4) was detd. Isotype I strains (with a 1.8-kb *tbpB* gene) was less represented than isotype II strains (19.4 vs. 80.6%). Isotype I was restricted to the ET-37 complex strains, while isotype II was found in all four clonal complexes. The extent of the allelic diversity of *tbpB* in these two groups was studied by PCR restriction anal. and sequencing of 10 new *tbpB* genes. Four major *tbpB* gene variants were characterized: B16B6 (representative of isotype I) and M982, B283,

**THIS PAGE BLANK (USPTO)**



and 8680 (representative of isotype II). The relevance of these variants was assessed at the antigenic level by the detn. of cross-bactericidal activity of purified IgG preps. raised to the corresponding recombinant TbpB (rTbpB) protein against a panel of 27 strains (5 of isotype I and 22 of isotype II). The results indicated that rTbpB corresponding to each variant was able to induce cross-bactericidal antibodies. However, the no. of strains killed with an anti-rTbpB serum was slightly lower than that obtained with an anti-TbpA+B complex. None of the sera tested raised against an isotype I strain was able to kill an isotype II strain and vice versa. None of the specific antisera tested (anti-rTbpB or anti-TbpA+B complex) was able to kill all of the 22 isotype II strains tested. Moreover, using sera raised against the C-terminus domain of TbpB M982 (amino acids 352 to 691) or BZ83 (amino acids 329 to 669) fused to the maltose-binding protein, cross-bactericidal activity was detected against 12 and 7 isotype II strains, resp., of the 22 tested. These results suggest surface accessibility of the C-terminal end of TbpB. Altogether, these results show that although more than one rTbpB will be required in the compn. of a TbpB-based vaccine to achieve a fully cross-bactericidal activity, rTbpB and its C terminus were able by themselves to induce cross-bactericidal antibodies.

REFERENCE COUNT: 46 THERE ARE 46 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L20 ANSWER 15 OF 25 EMBASE COPYRIGHT 2003 ELSEVIER SCI. B.V.

ACCESSION NUMBER: 2000097426 EMBASE

TITLE: A furtive **pathogen** revealed.

AUTHOR: **Nassif X.**

CORPORATE SOURCE: X. Nassif, INSERM Unit 411, Fac. de Med. Necker-Enfants Naiades, 156 Rue de Vaugirard, 75015 Paris, France. [nassif@necker.fr](mailto:nassif@necker.fr)

SOURCE: Science, (10 Mar 2000) 287/5459 (1767-1768).

Refs: 7

ISSN: 0036-8075 CODEN: SCIEAS

COUNTRY: United States

DOCUMENT TYPE: Journal; (Short Survey)

FILE SEGMENT: 004 Microbiology  
008 Neurology and Neurosurgery  
017 Public Health, Social Medicine and Epidemiology  
026 Immunology, Serology and Transplantation  
037 Drug Literature Index

LANGUAGE: English

L20 ANSWER 16 OF 25 HCAPLUS COPYRIGHT 2003 ACS DUPLICATE 7

ACCESSION NUMBER: 1999:709902 HCAPLUS

DOCUMENT NUMBER: 132:59905

TITLE: Identification of regions of the chromosome of *Neisseria meningitidis* and *Neisseria gonorrhoeae* which are specific to the pathogenic *Neisseria* species

AUTHOR(S): **Perrin, Agnes; Nassif, Xavier;**

**Tinsley, Colin**

CORPORATE SOURCE: Laboratoire de Microbiologie, INSERM U411, Faculte de Medecine Necker-Enfants Malades, Paris, 75015, Fr.

**THIS PAGE BLANK (USPTO)**

09/830433

SOURCE: Infection and Immunity (1999), 67(11), 6119-6129  
CODEN: INFIBR; ISSN: 0019-9567  
PUBLISHER: American Society for Microbiology  
DOCUMENT TYPE: Journal  
LANGUAGE: English

AB *Neisseria meningitidis* and *Neisseria gonorrhoeae* give rise to dramatically different diseases. Their interactions with the host, however, do share common characteristics: they are both human pathogens which do not survive in the environment and which colonize and invade mucosa at their port of entry. It is therefore likely that they have common properties that might not be found in nonpathogenic bacteria belonging to the same genetically related group, such as *Neisseria lactamica*. Their common properties may be detd. by chromosomal regions found only in the pathogenic *Neisseria* species. To address this issue, we used a previously described technique (C. R. Tinsley and X. Nassif, Proc. Natl. Acad. Sci. USA 93:11109-11114, 1996) to identify sequences of DNA specific for pathogenic *neisseriae* and not found in *N. lactamica*. Sequences present in *N. lactamica* were phys. subtracted from the *N. meningitidis* Z2491 sequence and also from the *N. gonorrhoeae* FA1090 sequence. The clones obtained from each subtraction were tested by Southern blotting for their reactivity with the three species, and only those which reacted with both *N. meningitidis* and *N. gonorrhoeae* (i.e., not specific to either one of the pathogens) were further investigated. In a first step, these clones were mapped onto the chromosomes of both *N. meningitidis* and *N. gonorrhoeae*. The majority of the clones were arranged in clusters extending up to 10 kb, suggesting the presence of chromosomal regions common to *N. meningitidis* and *N. gonorrhoeae* which distinguish these pathogens from the commensal *N. lactamica*. The sequences surrounding these clones were detd. from the *N. meningitidis* genome-sequencing project. Several clones corresponded to previously described factors required for colonization and survival at the port of entry, such as IgA protease and PilC. Others were homologous to virulence-assocd. proteins in other bacteria, demonstrating that the subtractive clones are capable of pinpointing chromosomal regions shared by *N. meningitidis* and *N. gonorrhoeae* which are involved in common aspects of the host interaction of both pathogens.

REFERENCE COUNT: 51 THERE ARE 51 CITED REFERENCES AVAILABLE  
FOR THIS RECORD. ALL CITATIONS AVAILABLE  
IN THE RE FORMAT

L20 ANSWER 17 OF 25 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.  
ACCESSION NUMBER: 1999:457991 BIOSIS  
DOCUMENT NUMBER: PREV199900457991  
TITLE: Meningococcal vaccine: Myth and reality.  
AUTHOR(S): Nassif, X. (1)  
CORPORATE SOURCE: (1) faculte Necker-Enfants-Malades, Inserm U 411,  
156, rue de Vaugirard, 75015, Paris France  
SOURCE: Archives de Pediatrie, (1999) Vol. 6, No. SUPPL. 3,  
pp. 647-649.  
ISSN: 0929-693X.  
DOCUMENT TYPE: Article  
LANGUAGE: French

L20 ANSWER 18 OF 25 HCAPLUS COPYRIGHT 2003 ACS DUPLICATE 8  
ACCESSION NUMBER: 1998:71228 HCAPLUS  
DOCUMENT NUMBER: 128:164910

Searcher : Shears 308-4994

**THIS PAGE BLANK (USPTO)**

09/830433

TITLE: Genes and gene products specific to  
**pathogenicity of Neisseria**  
meningitidis, methods for obtaining them and  
their biological applications

INVENTOR(S): **Nassif, Xavier; Tinsley, Colin**  
; Achtman, Mark; Ruelle, Jean-Louis; Vinals,  
Carla; Merker, Petra

PATENT ASSIGNEE(S): Institut National De La Sante Et De La Recherche  
Medicale (INSERM), Fr.; Max-Planck-Gesellschaft  
Zur Forderung Der Wissenschaften E.V., Berlin;  
Smithkline Beecham; Nassif, Xavier; Tinsley,  
Colin; Achtman, Mark; Ruelle, Jean-Louis;  
Vinals, Carla; Merker, Petra

SOURCE: PCT Int. Appl., 150 pp.  
CODEN: PIXXD2

DOCUMENT TYPE: Patent

LANGUAGE: French

FAMILY ACC. NUM. COUNT: 1

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9802547	A2	19980122	WO 1997-FR1295	19970711
WO 9802547	A3	19980409		
W:		AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM		
RW:		GH, KE, LS, MW, SD, SZ, UG, ZW, AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG		
FR 2751000	A1	19980116	FR 1996-8768	19960712
FR 2751000	B1	19981030		
AU 9736977	A1	19980209	AU 1997-36977	19970711
AU 730423	B2	20010308		
EP 951552	A2	19991027	EP 1997-933727	19970711
R:		AT, BE, CH, DE, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI		
JP 2001504684	T2	20010410	JP 1998-505685	19970711
US 2002164603	A1	20021107	US 2001-928457	20010814

PRIORITY APPLN. INFO.: FR 1996-8768 A 19960712  
WO 1997-FR1295 W 19970711  
US 1999-214759 B1 19990422

AB DNA sequences that are found in **Neisseria meningitidis**  
that are unique to it, specific to **pathogenesis**, and not  
found in *N. gonorrhoeae*, *N. lactamica* or *N. cinerea* are cloned by  
representational difference anal. A no. of genes assocd. with  
**pathogenesis** that are found in *N. meningitidis* and  
*N. gonorrhoeae* including the genes of biosynthesis of the  
polysaccharide capsule (*frpA*, *frpC*, *porA*), *pilC*, the genes for  
rotamase, IgA protease, pilin, transferring-binding proteins and  
opacity proteins and the sequence IS1106. The genes map in clusters  
in three regions of the chromosome. The gene products can be used  
as antigens in the raising of antibodies for diagnostic or  
therapeutic uses, e.g. specific immunoassays or **vaccines**.  
The roles of the genes in **pathogenesis** can be studied by

**THIS PAGE BLANK (USPTO)**

09/830433

targeted deletion.

L20 ANSWER 19 OF 25 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.  
DUPLICATE 9

ACCESSION NUMBER: 1998:377538 BIOSIS  
DOCUMENT NUMBER: PREV199800377538  
TITLE: Transferrin-binding proteins as candidates for a  
broadly cross-reactive vaccine against serogroup B  
meningococcal disease.  
AUTHOR(S): Quentin-Millet, M. J.; **Rokbi, B.**; Mignon,  
M.; Maitre-Wilmotte, G.; Danve, B.; **Renauld,**  
**G.**; Lissolo, L.  
CORPORATE SOURCE: Pasteur Merieux Connaught, 1541 avenue Marcel  
Merieux, 69280 Marcy l'Etoile Cedex France  
SOURCE: Biotechnologia Aplicada, (Jan.-March, 1998) Vol. 15,  
No. 1, pp. 41.  
Meeting Info.: Selected Papers from Congreso  
Biotechnologia Habana '97 Havana, Cuba December 1-6,  
1997  
ISSN: 0864-4551.  
DOCUMENT TYPE: Conference  
LANGUAGE: English

L20 ANSWER 20 OF 25 HCAPLUS COPYRIGHT 2003 ACS DUPLICATE 10

ACCESSION NUMBER: 1997:491635 HCAPLUS  
DOCUMENT NUMBER: 127:94113  
TITLE: Preparation of libraries of antibody genes in  
expression vectors capable of undergoing  
recombination for generation of new antibodies  
INVENTOR(S): Sodoyer, Regis; **Aujame, Luc**; Geoffroy,  
Frederique; **Bouchardon, Annabelle**  
PATENT ASSIGNEE(S): Pasteur Merieux Serums & Vaccins, Fr.; Sodoyer,  
Regis; Aujame, Luc; Geoffroy, Frederique;  
Bouchardon, Annabelle  
SOURCE: PCT Int. Appl., 37 pp.  
CODEN: PIXXD2  
DOCUMENT TYPE: Patent  
LANGUAGE: French  
FAMILY ACC. NUM. COUNT: 1  
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9720923	A1	19970612	WO 1996-FR1938	19961204
W: CA, JP, US				
RW: AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE				
FR 2741892	A1	19970606	FR 1995-14325	19951204
FR 2741892	B1	19980213		
CA 2239490	AA	19970612	CA 1996-2239490	19961204
EP 865487	A1	19980923	EP 1996-941701	19961204
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI				
US 6174708	B1	20010116	US 1999-101629	19990122
PRIORITY APPLN. INFO.: FR 1995-14325 A 19951204				
WO 1996-FR1938 W 19961204				

AB Libraries of genes for antibody light and heavy chains, including  
const. and variable regions, in bacteriophage or phagemid vectors

Searcher : Shears 308-4994

**THIS PAGE BLANK (USPTO)**



09/830433

that are capable of undergoing in vivo recombination are described for use in the generation of new antibody types in bacterial hosts. One suitable mechanism is xis-mediated excision. Sep. vectors are used for light and heavy chain genes and the variable regions are manufd. as fusion proteins with a capsid protein to direct surface display. The vectors can undergo one or more irreversible recombinations to generate recombinant final vectors with possible new genes for variable regions. The method increases the efficiency of generation and detection of new antibody types.

L20 ANSWER 21 OF 25 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.  
DUPLICATE 11

ACCESSION NUMBER: 1998:144177 BIOSIS  
DOCUMENT NUMBER: PREV199800144177  
TITLE: What do we know about the entry of Neisseria meningitidis into the meninges.  
AUTHOR(S): Nassif, X. (1); Pujol, C.; Tinsley, C.; Morand, P.; Eugene, E.; Marceau, M.; Perrin, A.; Pron, B.; Taha, M.-K.  
CORPORATE SOURCE: (1) INSERM U411, Lab. Microbiologie, Fac. Med. Necker-Enfants Malades, 156 Rue de Vaugirard, 75015 Paris France  
SOURCE: Bulletin de l'Institut Pasteur, (Oct.-Dec., 1997) Vol. 95, No. 4, pp. 219-235. ISSN: 0020-2452.  
DOCUMENT TYPE: General Review  
LANGUAGE: English

L20 ANSWER 22 OF 25 HCAPLUS COPYRIGHT 2003 ACS

ACCESSION NUMBER: 1996:386028 HCAPLUS  
DOCUMENT NUMBER: 125:56214  
TITLE: Hemoglobin receptors from *Neisseriae* meningitidis and *N. gonorrhoeae*, cDNA sequences, and vaccine development  
INVENTOR(S): Stojiljkovic, Igor; So, Magdalene; Hwa, Vivian; Heffron, Fred; Nassif, Xavier  
PATENT ASSIGNEE(S): Oregon Health Sciences University, USA  
SOURCE: PCT Int. Appl., 103 pp. CODEN: PIXXD2  
DOCUMENT TYPE: Patent  
LANGUAGE: English  
FAMILY ACC. NUM. COUNT: 2  
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9612020	A2	19960425	WO 1995-US13623	19951017
WO 9612020	A3	19960523		
W:	AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LT, LU, LV, MD, MG, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TT			
RW:	KE, MW, SD, SZ, UG, AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG			
US 5698438	A	19971216	US 1994-326670	19941018
US 6121037	A	20000919	US 1995-537361	19951002
AU 9540073	A1	19960506	AU 1995-40073	19951017

Searcher : Shears 308-4994

**THIS PAGE BLANK (USPTO)**

09/830433

AU 705509            B2    19990527  
EP 793720            A2    19970910            EP 1995-938842    19951017  
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LI, LU, MC, NL,  
PT, SE  
JP 10508469           T2    19980825            JP 1995-513482    19951017  
NO 9701768            A    19970603            NO 1997-1768      19970417  
FI 9701634            A    19970616            FI 1997-1634      19970417  
US 6277382            B1    20010821            US 1997-817707    19970819  
PRIORITY APPLN. INFO.:            US 1994-326670    A2 19941018  
   US 1995-537361      19951002  
   WO 1995-US13623    W 19951017

AB    The present invention relates to novel bacterial Hb receptor proteins and genes that encode such proteins. The invention is directed toward the isolation, characterization, diagnostic and therapeutic use of bacterial Hb receptor proteins, nucleic acids encoding such proteins, recombinant expression constructs comprising such nucleic acids and cells transformed therewith, and antibodies and epitopes of such Hb receptor proteins. The invention relates particularly to Hb receptor proteins and genes encoding such proteins from **Neisseria** species, esp. *N. meningitidis* and serotypes thereof, and *N. gonorrhoeae*. Methods for the diagnostic and therapeutic use of the proteins, epitopes, antibodies and nucleic acids of the invention are also provided, including the use of the proteins, epitopes, antibodies and nucleic acids of the invention for the prodn. of **vaccines** effective in providing **immunization** of a human against infection by **pathogenic** bacteria of **Neisseria** species.

L20 ANSWER 23 OF 25 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.

ACCESSION NUMBER: 1997:275422 BIOSIS

DOCUMENT NUMBER: PREV199799574625

TITLE: Bacterial meningitis: Treatment and prevention strategies.

AUTHOR(S): Bingen, F.; Bourillon, A.; Clavaud, R.; Geslin, P.; Gicquel, B.; Guerin, N.; Livartowski, A.;

SOURCE: **Nassif, X.**; Reinert, P.; Riou, J.-Y.; et al.  
Bingen, E.; Bourillon, A.; Clavaud, R.; Geslin, P.; Gicquel, B.; Guerin, N.; Livartowski, A.; **Nassif, X.**; Reinert, P.; INSERM. (1996) pp. xi+167p. Bacterial meningitis: Treatment and prevention strategies. Meningites bacteriennes: Strategies de traitement et de prevention. Publisher: INSERM (Institut National de la Sante et de la Recherche Medicale) 101, rue de Tolbiac, 75654 Paris Cedex 13, France. ISBN: 2-85598-686-9.

DOCUMENT TYPE: Book

LANGUAGE: French

AB    This monograph on treatment and prevention strategies for bacterial meningitis is based on approximately 500 articles and represents the collective work of a group of experts assembled by the INSERM, the French National Institute of Health and Medical Research. The book provides a useful reference to health care professionals, medical instructors and researchers, public health professionals, epidemiologists and those involved with health care policy. The first of the book's two sections represents a detailed analysis of the current world literature on various aspects of bacterial meningitis that was developed during a series of meetings of the

**THIS PAGE BLANK (USPTO)**

expert group. The first part of this analytical section is devoted to general aspects of bacterial meningitis and contains chapters on clinical data and pathophysiology, germs and diagnostic methods, epidemiological and socioeconomic data. The following part discusses treatment and **vaccinal** prevention of bacterial meningitis, including antibiotic treatment and prophylaxis, current **vaccines** and **vaccine** strategies for dealing with four of the bacterial meningitis **pathogens** (*Haemophilus influenzae*- b, *Neisseria meningitidis*, *Streptococcus pneumoniae* and *Mycobacterium tuberculosis*). The third part of the analytical section focuses on prospects for **vaccines** against meningitis, especially **vaccines** against the *Neisseria meningitidis*, *Streptococcus pneumoniae* and *Mycobacterium tuberculosis* **pathogens**. The second section of the book provides a synthesis of the information available on bacterial meningitis, its treatment and the research and development of **vaccines** for its prevention. This section concludes with seven recommendations developed by the group of experts for the therapeutic management and prevention of bacterial meningitis. Tables, graphs, diagrams and end-of-chapter bibliographies supplement the text.

L20 ANSWER 24 OF 25 MEDLINE

ACCESSION NUMBER: 96009214 MEDLINE

DOCUMENT NUMBER: 96009214 PubMed ID: 7553571

TITLE: Interaction of **pathogenic neisseriae** with nonphagocytic cells.

AUTHOR: Nassif X; So M

CORPORATE SOURCE: Institut National de la Sante et de la Recherche Medicale U411, Faculte de Medecine Necker-Enfants Malades, Universite Rene Descartes, Paris, France.

SOURCE: CLINICAL MICROBIOLOGY REVIEWS, (1995 Jul) 8 (3) 376-88. Ref: 148  
Journal code: 8807282. ISSN: 0893-8512.

PUB. COUNTRY: United States

DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)  
General Review; (REVIEW)  
(REVIEW, TUTORIAL)

LANGUAGE: English

FILE SEGMENT: Priority Journals

ENTRY MONTH: 199510

ENTRY DATE: Entered STN: 19951227  
Last Updated on STN: 19951227  
Entered Medline: 19951025

AB The ability to interact with nonphagocytic cells is a crucial virulence attribute of the meningococcus and the gonococcus. Like most bacterial **pathogens**, *Neisseria meningitidis* and *Neisseria gonorrhoeae* initiate infections by colonizing the mucosal epithelium, which serves as the site of entry. After this step, both bacteria cross the intact mucosal barrier. While *N. gonorrhoeae* is likely to remain in the subepithelial matrix, where it initiates an intense inflammatory reaction, *N. meningitidis* enters the bloodstream, and eventually the cerebrospinal fluid to cause meningitis. Both **pathogens** have evolved very similar mechanisms for interacting with host cells. Surface structures that influence bacterium-host interactions include pili, the meningococcal class 5 outer membrane proteins or the gonococcal opacity proteins, lipooligosaccharide, and the

**THIS PAGE BLANK (USPTO)**

09/830433

meningococcal capsule. This review examines what is known about the roles these structures play in bacterial adhesion and invasion, with special emphasis, on pilus-mediated adhesion. Finally, the importance of these structures in **neisserial pathogenesis** is discussed.

L20 ANSWER 25 OF 25 HCAPLUS COPYRIGHT 2003 ACS DUPLICATE 12  
ACCESSION NUMBER: 1993:20647 HCAPLUS  
DOCUMENT NUMBER: 118:20647  
TITLE: Antibodies recognizing a variety of different structural motifs on meningococcal Lip antigen fail to demonstrate bactericidal activity  
AUTHOR(S): **Tinsley, Colin R.**; Virji, Mumtaz; Heckels, John E.  
CORPORATE SOURCE: Med. Sch., Southampton Univ., Southampton, SO9 4XY, UK  
SOURCE: Journal of General Microbiology (1992), 138(11), 2321-8  
CODEN: JGMIAN; ISSN: 0022-1287  
DOCUMENT TYPE: Journal  
LANGUAGE: English  
AB The **neisserial** Lip antigen is a conserved antigen assocd. with the **pathogenic Neisseria** species, and is composed of multiple repeats of a consensus pentapeptide. A series of monoclonal antibodies reacting with meningococcal Lip antigen were subjected to epitope mapping, using solid-phase synthetic peptides based on the concensus repeat sequence. The antibodies were found to recognize different continuous epitopes based on the consensus sequence. One monoclonal antibody was utilized in affinity chromatog. to obtain purified Lip antigen and the antigen was used for **immunization** of mice. The resulting antisera did not recognize Lip antigen on Western blots but reacted specifically with Lip antigen in immune pptn. expt., indicating that the predominant polyclonal immune response was directed against conformational epitopes. Despite the diversity of both continuous and conformational epitopes recognized by the antibodies produced, none of the antibodies demonstrated the ability to promote complement-mediated bactericidal activity. Thus, despite its initial apparent promise as a potential **vaccine** candidate the case for the inclusion of Lip antigen in **vaccine** formulation cannot be supported at present.

FILE 'HOME' ENTERED AT 10:19:09 ON 28 JAN 2003

**THIS PAGE BLANK (USPTO)**